1990

On the stability of linear stochastic difference equations

Patrick René Homblé
Iowa State University

INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

U·M·I

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313 761-4700 800 521-0600

Order Number 9101353

On the stability of linear stochastic difference equations

Homblé, Patrick René, Ph.D.

Iowa State University, 1990

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

On the stability of linear stochastic difference equations

by

Patrick René Homblé

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major : Statistics

Approved:

Members of the Committee:

Signature was redacted for privacy.

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University

Ames, Iowa

1990

TABLE OF CONTENTS

NOTATION

$\ll$	absolutely continuous with respect to.
$\cong$	isomorphic to.
$\mathbf{1}$	the identity map.
$\|x\|$	the Euclidean norm of $x \in \mathbb{R}^d$.
$\|\|A\|\|$	a norm (to be specified) for the operator A.
$\setminus$	set difference.
$\sim$	"distributed as" for random variables.
$\bigcup, \bigcap$	$\bigcup \{A_\lambda ; \lambda \in \Lambda\}$ is the union of the A_λ's over $\lambda \in \Lambda$ and similarly for the intersection $\bigcap$.
$\overline{A}$	the closure of A.
A^c	the complement of A.
$\mathcal{B}(E)$	the σ-algebra of Borel subsets of E.
$B(E)$	the bounded measurable functions on the space E.
C	a (maximal) invariant control set (p. 81).
$\mathbb{C}$	the complex numbers.
C^1	differentiable.
C^∞	infinitely differentiable.
$C(E)$	the continuous maps from E to E.
$C(t, m)$	a (multiplicative) cocycle associated with a flow (deterministic case) (p. 81).
$C(t, x, m)$	a (multiplicative) cocycle associated with a flow (stochastic case) (p. 82).

$C(n, \omega)$	the cocycle associated with the flow of $x_{n+1} = A(\xi_n(\omega))x_n$ (p. 186).
C_x^k	a generalized controllability matrix (p. 147).
$\mathscr{D}(A)$	the domain of the operator A.
d_i	the multiplicity of the (forward) Lyapunov exponent λ_i (p. 64).
∂A	the boundary of A.
det A	the determinant of A.
DF	the Jacobian matrix of F (p. 9).
Diff (M)	the set of all diffeomorphisms from M to M.
dim (E)	the dimension of E.
$\mathscr{E}$	a Banach space.
$E^{\mathbb{N}}$	the path space of an E valued process.
E_i	$L_i \cap L^-_{p+1-i}$ (p. 67).
$\mathscr{F}$	a σ-algebra.
F_*	the differential of F (p. 17).
$F(n, x, B)$	$P(T(x, B) = n)$ (p. 48).
$\gamma(p)$	$g(p) / p$ (p. 264).
$\mathcal{G}$	an arbitrary group in Diff (M).
$\mathcal{G}(\mathcal{X})$	the group generated by $\mathcal{X} \subset M$.
$G(x, B)$	$\sum_{n=1}^{\infty} \mu(n, x, B)$ (p. 48).
$Gl(d, \mathbb{R})$	the Lie group of d × d real valued invertible matrices.
$gl(d, \mathbb{R})$	the Lie group of d × d real valued matrices.
g (p)	the p^{th} moment Lyapunov exponent (pp. 204, 206).
Gx	the orbit generated by $\mathcal{G}$ applied to $x \in M$ (p. 27).

H_t	shift operator on a path space (H_t A (s) = A (s+t)) (pp. 68, 185).
I (r)	level–1 entropy function (p. 216).
B	the indicator function of the set B.
int A	the interior of A.
λ	the maximal (forward) Lyapunov exponent (= λ_1).
λ_i	the i^{th} (forward) Lyapunov exponent (p. 64).
λ (x)	the (forward) Lyapunov exponent associated with the initial value x (p. 63).
$\lambda(\omega, x)$	the (forward) Lyapunov exponent associated with the initial value x (stochastic case) (p. 187).
L_i	the i^{th} subspace in the filtration associated with a (forward) Lyapunov spectrum (p. 64).
L (x, B)	$P\left[\bigcup_{n=1}^{\infty}\left[X_n \in B\right] \,\middle|\, X_0 = x\right]$ (p. 48).
$\mu(.,.)$	the kernel of a Markov processes (usually the $\{\xi_n\}$ process) (pp. 39, 42).
$\bar{\mu}(.,.)$	the kernel of the Markov pair process $\{(x_n, \xi_n)\}$.
μ (n, x, B)	the n^{th} step transition probability from x to B.
M	a connected C^∞ Riemannian manifold (p. 15).
M^d	a connected C^∞ Riemannian manifold of dimension d (p. 15).
M (E)	the measurable functions on E.
$\mathcal{M}(E)$	the set of all measures on the measurable space $(E, \mathcal{B}(E))$.
m_C	the volume element m_M restricted to the maximal invariant control set C.
m_M	the volume element on the manifold M.

$\mathbb{N}$	the natural numbers.
$\mathbb{N}_0$	$\mathbb{N} \setminus \{0\}$.
π	the unique invariant probability measure for the pair process $\{(x_n, \xi_n)\}$.
π_ξ	the unique invariant probability measure for the noise process $\{\xi_n\}$.
p	the number of distinct (forward) Lyapunov exponents (p. 64).
$\mathbb{P}^{d-1}$	the projective space in $\mathbb{R}^d$ (p. 11).
$\mathcal{P}(E)$	the set of all probability measures on the measurable space $(E, \mathcal{B}(E))$.
Q	the support of the measure π_ξ.
$\mathbb{Q}$	the rational numbers.
$\mathbb{Q}_0$	$\mathbb{Q} \setminus \{0\}$.
$Q(x, B)$	$P(X_n \in B \text{ i.o.} \mid X_0 = x)$ (p. 48).
ρ	a metric.
$\rho(x, y)$	the distance between x and y in the metric ρ.
$\mathbb{R}^d$	the d dimensional Euclidean space.
$\mathbb{R}_0^d$	$\mathbb{R}^d \setminus \{0\}$.
$R(x, B)$	$\sum_{n=1}^{\infty} n\, F(n, x, B)$.
Σ	a dynamical control system (p. 29).
$\mathbb{S}^{d-1}$	the sphere in $\mathbb{R}^d$ (p. 11).
$\mathcal{S}$	an arbitrary semigroup in Diff (M).
$\mathcal{S}(\mathcal{U})$	the semigroup generated by $\mathcal{U} \subset M$.
Sx	the (positive) orbit generated by $\mathcal{S}$ applied to $x \in M$ (p. 28).

$S^{-}x$	the negative orbit generated by S^{-1} applied to $x \in M$ (p. 100).
$S^{n}_{x\,\xi}$	the value at time n of the stochastic orbit of x, starting with $\xi_0(\omega) = \xi$ (p. 147).
$S_{x\,\xi}$	the stochastic orbit of x, starting at $\xi_0(\omega) = \xi$ $(= \bigcup \{S^{n}_{n\,\xi}\,;\, n \geq 1\})$ (p. 147).
supp (ν)	the support of the measure ν.
T	a time set ($T = \mathbb{N}$ or $\mathbb{R}^{+}$).
$T(M)$	the tangent bundle of the manifold M (p. 16).
$T_p(M)$	the tangent space to the manifold M at the point $p \in M$ (p. 16).
$\tau(x, B)$	$\inf(n > 0\,;\, X_n \in B \mid X_0 = x)$ (p. 48).
$\{T(t)\}$	a semigroup, $t \in T$ (p. 44).
$x(t, x_0)$	the solution at time $t \in T$ ($T = \mathbb{N}$ or $\mathbb{R}^{+}$) of a dynamical system with initial value $x(0) = x_0$.
$x(t, x_0, \omega)$	the solution at time $t \in T$ ($T = \mathbb{N}$ or $\mathbb{R}^{+}$) of a stochastic dynamical system with initial value $x(0) = x_0$.
$\mathbb{I}$	the integers.

1. INTRODUCTION

In dynamical system theory, processes evolving over time are often described by differential or difference equations of a deterministic nature. But such processes are almost always submitted to random perturbations arising from various sources like temperature or pressure variations, human responses, and, in general, from the unpredictable actions of the environment in which they evolve.

Therefore, the observed behavior of the systems under study often departs from the mathematical solution imposed by their deterministic description. Usually, such departures from theoretical behavior are corrected by feedback or outside (human) intervention. Nevertheless, the need for larger scale projects, better safety standards, optimization of the systems performance, and, in general, a more accurate description of the phenomenon under scrutiny has made it necessary to incorporate random effects in the equations used to model these processes.

The literature related to stochastic dynamical systems goes back at least to the thirties and has become quite abundant for both continuous and discrete time processes. These studies deal with various problems such as optimization, filtering, stability, parameter estimation, etc. Over the last five to ten years, a substantial amount of work on the stability of stochastic systems has been performed using Lyapunov exponents (called characteristic numbers by Lyapunov (1949)). Besides Lyapunov's work which goes back to the late nineteenth century, the foundations of this approach were set in publications dating back to the sixties. We should specially mention Oseledeč's (1968) for his well–known Multiplicative Ergodic Theorem,

Furstenberg (1963) as well as Furstenberg and Kesten (1960) for their work on products of random matrices, and finally Has'minskiĭ's (1967, 1980) for his pioneering work on the almost sure stability of linear stochastic systems and his unified overview of older results.

In several instances, the tools of deterministic control theory are used to obtain such stability results. The basic idea is to "replace" the random trajectories of the noise process by control functions in such a way that the associated deterministic control system so constructed does mimic the possible behaviors of the random system.

First found in Kliemann (1979) and then in Arnold and Kliemann (1983), the above approach was used by Arnold et al. (1986a) in their study of the continuous time stochastic system

$$\dot{x}(t) = A(\xi(t))\, x(t), \quad x(0) = x_0 \in \mathbb{R}^d_0,$$

where $A : M \longrightarrow gl(d, \mathbb{R})$ is an analytic mapping from an analytic connected Riemannian manifold M into the space $gl(d, \mathbb{R})$, the real valued $d \times d$ matrices, and where the noise $\{\xi(t)\}$ is a stationary ergodic diffusion process on M described by a stochastic differential equation. Using geometric control theory, these authors were able to show that, under some nondegeneracy assumption and via projection onto the system on the (compact) projective space $\mathbb{P}^{d-1}$, the pair process $\{(s_t, \xi_t)\}$, $s_t = \frac{x_t}{|x_t|}$, possesses a unique invariant probability measure. From there, they deduced the uniqueness (with probability one) of a nonrandom Lyapunov exponent, which is then used to study the asymptotic stability of the stochastic system.

This study was pursued in a series of closely related papers of which we mention only four. Arnold and Kliemann (1987) and Kliemann (1987) engaged in detailed investigations concerning the existence of a unique invariant probability measure, a crucial step in this approach. Arnold et al. (1986b) and Arnold and Kliemann (1986) pursued a finer study of the system's stability properties using the notion of moment Lyapunov exponents and large deviation theory.

A nice overview of the ideas found in the above references can be found in Kliemann (1988).

Discrete time stochastic systems have also been studied by various authors. In their book, Bougerol and Lacroix (1985) studied the system (on $\mathbb{R}_0^d$ and $\mathbb{P}^{d-1}$)

$$x_{n+1} = A_n x_n,$$

where $\{A_n\}$ is a sequence of iid real valued invertible matrices (i.e., $A_n \in Gl(d, \mathbb{R})$), via Furstenberg's (and others) theory on the product of random matrices. A stability study for such systems can be found in Bougerol (1987). Lyapunov exponents are widely used in these studies, but the approach taken is quite different from the one described above and, without the tools of geometric control theory, these authors did not obtain a unique exponent. Because of this, no further reference to this work will be made in this thesis. Note nevertheless that, in a subsequent paper, Bougerol (1988) extended these ideas to (among others) discrete systems of the form $x_{n+1} = A(\xi_n) x_n$, where $\{\xi_n\}$ is a Markov chain on a state space E and A is a map from E to $Gl(d, \mathbb{R})$, and further to more general continuous time setups referred to as multiplicative Markovian processes (see Bougerol (1985, 1986a, and 1986b)).

Finally, we should also mention the works of Meyn (1989) and Meyn and Caines (1988) who studied the nonlinear discrete stochastic system on $\mathbb{R}^d$

$$x_{n+1} = f(x_n, \xi_n),$$

where $f : \mathbb{R}^d \times \mathbb{R}^n \longrightarrow \mathbb{R}^d$ is a smooth map and $\{\xi_n\}$ is a sequence of iid random variables. These authors used deterministic control theory to obtain conditions which guarantee the existence of an invariant probability measure for the Markov chain $\{x_n\}$, but did not discuss asymptotic stability properties via the theory of Lyapunov exponents.

The purpose of this work is to "combine" all the above approaches. Namely, we will study the discrete time stochastic system on $\mathbb{R}_0^d$

$$x_{n+1} = A(\xi_n)\, x_n,$$

where (with further specifications given in the text) $\{\xi_n\}$ is a stationary ergodic Markov chain taking values on a connected C^∞ Riemannian manifold W and A is a mapping from W into $Gl(d, \mathbb{R})$.

This investigation is modeled according to the work (in the continuous time case) of Arnold et al. (1986a), i.e., via the tools of geometric control theory applied to the projection onto $\mathbb{P}^{d-1}$ of the above stochastic system, $s_{n+1} = f(s_n, \xi_n)$, where $s_n = x_n\, |x_n|^{-1}$ and $f(s, \xi) = |A(\xi)\, s|^{-1} A(\xi)\, s$.

The two main results found in this thesis are:

1) Under some assumptions to be specified, the Markov chain $\{(s_n, \xi_n)\}$ possesses a unique invariant (probability) measure. This is Theorem 4.2.1.

2) Using the existence of this invariant probability measure, we show that the Lyapunov exponent associated with the stochastic system $x_{n+1} = A(\xi_n) x_n$ is almost surely unique and independent of the initial value. This is Theorem 5.2.1.

Theorem 4.2.1 is a discrete analog to Theorem 5.1 found in Arnold and Kliemann (1987). While these authors used a Lie algebra condition to prove this result, we use a lower semi–continuity assumption on the density of the one–step transition probabilities of the noise $\{\xi_n\}$ combined with the assumption of weak stochastic controllability as defined in Meyn and Caines (1988). In fact, under the lower semi–continuity assumption, the weak stochastic controllability assumption is shown to be itself implied by a Lie algebra condition (see Subsection 4.1: Theorem 1, Proposition 1, and the discussion thereafter).

Theorem 5.2.1 is then used to initiate a study of the asymptotic stability behavior of the system, using both almost sure Lyapunov and moment Lyapunov exponents. This theorem and the other results in Section 5 are discrete time analogs of several results found in Arnold et al. (1986b) and Arnold and Kliemann (1986).

Finally, to illustrate the above, a computer simulation is performed on a discretized version of the linear oscillator with damping and restoring force.

The thesis is organized in five main sections. Section 2 consists of a review of most of the notions used in the text. Because this work involves concepts from different areas of mathematics, stochastic processes theory (Markov chains), and control system theory, this review was intentionally made very verbose. It is

expected that the reader will skip over the topics with which he/she is familiar.

Section 3 is devoted to the study of discrete deterministic systems via their associated semigroup. The notions of orbit under a semigroup action, of maximal invariant control set, and of accessibility, controllability, and transitivity are defined and carefully studied for two main reasons:

1) Some of these notions will play a crucial role in the remainder of the thesis, especially in establishing the existence of a unique invariant measure for the pair process $\{(s_n, \xi_n)\}$.
2) There is a general interest in comparing the properties of these entities in the discrete time case with their properties in the continuous time case.

In Section 3, much emphasis was put on the notion of semigroup but all the relevant examples and counterexamples are given using semigroups which could arise from controlled difference equations.

Section 4 is central to the thesis. There we state and discuss a series of assumptions which, together with the results of Section 3, are used to establish the key result of the existence of a unique invariant probability measure for the pair process $\{(s_n, \xi_n)\}$.

In Section 5, we use this invariant measure to prove the existence of a unique (with probability one) Lyapunov exponent and proceed to investigate the stability properties of our original stochastic system. We do this first via this unique

Lyapunov exponent. Next, we carry on with a finer study using the notion of moment Lyapunov exponents. Some large deviation results are also given.

Section 6 is then devoted to the computer simulation mentioned above. At the beginning of the chapter, we describe the discretized version of the random linear oscillator on which the simulation is based and show that the assumptions specified in Sections 4 and 5 are satisfied. The end of Section 6 is devoted to the simulation results. Excerpts of the Pascal program used are given in the Appendix.

2. MATHEMATICAL REVIEW

2.1. Mappings

We first give the definition of some terms frequently used in the text, recall some results, and set up some notation. More details concerning this material can be found in any standard book on topology, e.g., Munkres (1975), or on differential geometry, e.g., Boothby (1986) and Helgason (1978).

Let $(E_1, \mathcal{T}_1)$ and $(E_2, \mathcal{T}_2)$ be two topological spaces.

Definition 1.

A mapping $f : E_1 \longrightarrow E_2$ is said to be a homeomorphism if f is a bijection such that both f and f^{-1} are continuous.

If E_1 and E_2 are differentiable manifolds, the mapping f will be called a diffeomorphism if it is a homeomorphism such that both f and f^{-1} are C^∞, i.e., infinitely often differentiable (see Subsection 2.3 for relevant definitions).

Note that, by definition, if f is a homeomorphism, $U \in \mathcal{T}_1$, and $V \in \mathcal{T}_2$, then $f(U) \in \mathcal{T}_2$ and $f^{-1}(V) \in \mathcal{T}_1$, i.e., $f(U)$ and $f^{-1}(V)$ are open. Moreover, when working on manifolds, a much deeper result is that, by Brouwer's theorem on invariance of domain, we then necessarily have $\dim(E_1) = \dim(E_2)$.

Definition 2.

Let $F : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ be a mapping defined by $F(x) = (f_1(x), \ldots, f_n(x))$, where $x \in \mathbb{R}^m$ is $(x_1, \ldots, x_m)$. Then $f_i : \mathbb{R}^m \longrightarrow \mathbb{R}$, $i = 1, \ldots, n$, are the coordinate functions of F. Moreover, if U is a subset of $\mathbb{R}^m$ and F is differentiable on U, then the matrix DF

given by the expression

$$DF \equiv \frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_m)} \equiv \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_m} \end{bmatrix}$$

is defined at each point $a \in U$. This matrix is called the Jacobian matrix of F.

Proposition 1.

A necessary and sufficient condition for the C^∞ map F to be a diffeomorphism from U to F (U) is that it be one–to–one and DF be nonsingular at every point of U.

Proof.

See Corollary 2.6.7, p. 46 in Boothby (1986). ■

To illustrate the use of the above proposition, consider the function $h : \mathbb{R} \longrightarrow \mathbb{R}$ defined by

$$h(x) = \begin{cases} \exp\left[-x^{-2}\right] + x & x \neq 0 \\ x & x = 0. \end{cases}$$

In view of Proposition 1, h is clearly a diffeomorphism. It is a useful map because, for all n, its n-th derivative evaluated at zero is zero. This function will be used to construct examples and counterexamples in Section 3.

Let $A \subset E_1$. We will write $\overline{A}$, int A, A^c, and ∂A for the closure, the interior, the complement, and the boundary of A, respectively. Recall that:

a) $\text{int } A = \overline{A} \setminus \partial A = \bigcup \{O ; O = U \cap A, \ U \in \mathcal{T}_1\}$,

b) $\partial A = \overline{A} \cap \overline{A^c} = \{x \in X : O \cap A \neq \phi \text{ and } O \cap A^c \neq \phi \text{ for all } O \in \mathcal{T}_1 \text{ with } x \in O\}$,

c) $\partial A \cap \text{int } A = \phi$, and

d) $\text{int } A = X \setminus \overline{A^c}$.

Definition 3.

Let $f : E_1 \longrightarrow E_2$ be a surjective map. Then f is said to be a quotient map provided $U \in \mathcal{T}_2$ if and only if $f^{-1}(U) \in \mathcal{T}_1$.

The quotient topology induced by f on E_2 is precisely the unique topology that makes f a quotient map.

We say that f is an open map if $U \in \mathcal{T}_1$ implies $f(U) \in \mathcal{T}_2$.

Note that f being a quotient map immediately implies that f is continuous (this property is even stronger than continuity since, for a continuous map, it is possible to have $f^{-1}(U)$ open in E_1 while U is closed in E_2). Also, by definition, open and continuous surjective maps are necessarily quotient maps.

Given an equivalence relation $\sim$ on $(E_1, \mathcal{T}_1)$, denote by $[x] \equiv \{y \in E_1 : y \sim x\}$ the equivalence class of x, and, for $A \subset E_1$, define $[A] = \bigcup \{ [a] ; a \in A\}$. Write $E_1/_\sim$ for the set of equivalence classes on E_1 under $\sim$ and define the natural mapping $\pi : E_1 \longrightarrow E_1/_\sim$ by $\pi(x) = [x]$. Make $E_1/_\sim$ a topological space by giving it the quotient topology (which renders π continuous).

Definition 4.

With the above notation and topology, $E_1/_\sim$ is called the quotient space of E_1 relative to the equivalence relation $\sim$.

Let us give some concrete examples of quotient spaces, which will be needed later.

Example 1.

Take $E_1 = \mathbb{R}^d_o \equiv \mathbb{R}^d \setminus \{0\}$, $d \geq 2$, and define $\sim$ by $x \sim y$ if $x = t\,y$ for some $t \neq 0$. The equivalence classes can then be viewed as lines through the origin. The quotient space $E_1/_\sim$ is denoted by $\mathbb{P}^{d-1}$ and called the projective space. Note that $\mathbb{P}^{d-1}$ can be "identified" with a half sphere in $\mathbb{R}^d_o$. Moreover, for U open in $\mathbb{R}^d_o$, the mapping $\pi : \mathbb{R}^d_o \longrightarrow \mathbb{P}^{d-1}$ satisfies $\pi(U) = [U] = \bigcup \{g_t(U) ; t \neq 0\}$ where $g_t : \mathbb{R}^d_o \longrightarrow \mathbb{R}^d_o$ is defined by $g_t(x) = t\,x$. Since g_t is clearly a homeomorphism for $t \neq 0$, $\pi(U)$ is open in $\mathbb{P}^{d-1}$, i.e., π is an open map. Also note that if the equivalence relation $\sim$ is defined by $x \sim y$ if $x = t\,y$ for $t > 0$, $E_1/_\sim$ is denoted by $\mathbb{S}^{d-1}$ and can be identified with $\{x \in \mathbb{R}^d_o : |x| = 1\}$, the sphere in $\mathbb{R}^d_o$.

Remark 1.

Let G be a group acting on a set E_1 (see Subsection 2.2). Define the subgroup $G_x \equiv \{g \in G : g\,x = x\}$. Let the equivalence relation $\sim$ be defined by $g_1 \sim g_2$ if $g_2 = g_1\,g$ for some $g \in G_x$. We then have that $[g] = \{h \in G : h = g\,f \text{ for some } f \in G_x\}$ and the quotient space $\bigcup \{[g] ; g \in G\}$ is denoted by G / G_x. Its elements are called cosets of G. This notation will be used in the next subsection.

Now, if we write Diff (M) for the collection of all diffeomorphisms on some manifold M (see Subsection 2.3), it will be necessary to look at Diff (M) as a topological space. Hence, we need to define a topology on this space. One way to achieve this is to assign to Diff (M) the open–compact topology defined by

U is open in Diff (M) if $U = \{f \in \text{Diff}(M) : f(K) \subset V\} \equiv [K, V]$ for some compact set $K \subset M$ and some open set $V \subset M$.

It is a simple exercise to verify that the topology induced on the subset $\text{Gl}(d, \mathbb{R})$ of $\text{Diff}(\mathbb{R}^d_0)$ by the open–compact topology on $\text{Diff}(\mathbb{R}^d_0)$ is equivalent to the metric topology on $\text{Gl}(d, \mathbb{R})$ defined by the basis of ϵ-balls,

$$B_\epsilon(A_0) \equiv \{B \in \text{Gl}(d, \mathbb{R}) : \|B - A_0\| < \epsilon\},$$

where $\|.\|$ is some norm on $\text{Gl}(d, \mathbb{R}) \subset \text{gl}(d, \mathbb{R})$. (Note that $\text{Gl}(d, \mathbb{R})$ itself is not a linear space.)

2.2 Topological Groups

This section is devoted to a brief review of the concept of topological groups and of some results needed later. The reader will find more details on this subject in the references already given in Subsection 2.1.

Definition 1.

Let $(G, \mathcal{T})$ be a topological space and assume that G is also a group. Then G is a topological group if, for all g, g_1, and $g_2 \in G$, the mappings

$$(g_1, g_2) \longmapsto g_1 g_2 \text{ and}$$
$$g \longmapsto g^{-1}$$

are both continuous.

Definition 2.

Let G be a group and X be a set. Then G is said to act on X (on the left) if there is a mapping $\theta : G \times X \longrightarrow X$ satisfying the following two conditions:

1) if e is the identity element of G, then $\theta(e, x) = x$ for all $x \in X$, and

2) if $g_1, g_2 \in G$, then $\theta(g_1, \theta(g_2, x)) = \theta(g_1 g_2, x)$ for all $x \in X$.

When G is a topological group, X a topological space, and θ is continuous, the action is called continuous.

Remark 1.

In order to simplify the notation, $\theta(g, x)$ will routinely be written as $g\,x$ and the two conditions in the above definition become $e\,x = x$ and $(g_1 g_2)\,x = g_1\,(g_2\,x)$.

Definition 3.

Let G be a group acting on a set X and let $x \in X$. The isotropy group of G at x, denoted G_x, is the subgroup of G leaving x invariant, i.e., $G_x = \{g \in G : g\,x = x\}$.

Definition 4.

Let G be a group acting on a set X. We say that the action of G is transitive if, for all $x \in X$, $\{y \in X : y = g\,x \text{ for some } g \in G\} = X$.

Theorem 1.

Let G be a locally compact topological group with a countable basis and let X be a locally compact Hausdorff space. Let G_x be the isotropy group of G at $x \in X$. Then

a) G_x is closed in G and

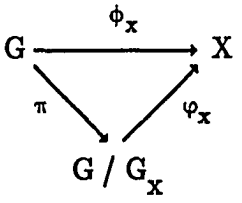
b) $\pi : G \longrightarrow G / G_x$ is an open and continuous map.

Moreover, if $\theta : G \times X \longrightarrow X$ is a transitive and continuous map and the mappings $\phi_x : G \longrightarrow X$ and $\varphi_x : G / G_x \longrightarrow X$ are defined by

$$\phi_x (g) = g\,x \quad \text{and} \quad \varphi_x ([g]) = [g]\,x,$$

we have

c) the diagram below is commutative with φ_x a homeomorphism, π open and continuous, and hence, ϕ_x open and continuous.



d) If, moreover, G is a Lie group (see Subsection 2.4) and θ is a C^∞ action, then φ_x is a diffeomorphism.

Proof.

a) See Theorem 2.3.2, p. 121 in Helgason (1978).

b) See Theorem 3.7.12, p. 94 in Boothby (1986)

c) See Theorem 2.3.2, p. 121 in Helgason (1978).

d) See Theorem 4.9.3, p. 167 in Boothby (1986).

■

2.3. Differentiable Manifolds

The purpose of this subsection is to review some basic notions in differential geometry. For more details, see, e.g., Boothby (1986) and Helgason (1978).

Definition 1.

A topological manifold of dimension d, say M (or M^d when the dimension is to be stressed), is a topological space with the following properties:

1) M is a Hausdorff space,
2) each point $p \in M$ has a neighborhood U_p which is homeomorphic to an open set V of $\mathbb{R}^d$, i.e., M is locally Euclidean, and
3) M has a countable basis of open sets.

It follows from this definition that M is locally connected, locally compact, the union of a countable collection of compact sets, normal, and metrizable (Theorem 1.3.6, p. 9 in Boothby (1986)). To avoid working on different components, we will from now on assume that all our manifolds (when used as state spaces) are connected. In particular, when we will use $\mathbb{R}^d_0$, it will be understood that $d \geq 2$.

Definition 2.

A coordinate neighborhood on a topological manifold M is a pair (U, φ) where U is an open set in M and φ is a homeomorphism between U and an open subset of $\mathbb{R}^d$.

Definition 3.

A C^∞ structure on a topological manifold M is a family $\mathcal{U} = \{(U_\gamma, \varphi_\gamma) ; \gamma \in \Gamma\}$ of coordinate neighborhoods such that

1) $\bigcup \{U_\gamma ; \gamma \in \Gamma\} = M$,
2) for any $\alpha, \gamma \in \Gamma$, the coordinate neighborhoods $(U_\alpha, \varphi_\alpha)$ and $(U_\gamma, \varphi_\gamma)$ are C^∞ compatible, i.e., $\varphi_\alpha \circ \varphi_\gamma^{-1}$ and $\varphi_\gamma \circ \varphi_\alpha^{-1}$ are diffeomorphisms on the open subsets $\varphi_\alpha(U_\alpha \cap U_\gamma)$ and $\varphi_\gamma(U_\alpha \cap U_\gamma)$ of $\mathbb{R}^d$, and

3) any coordinate neighborhood (V, ψ) compatible with every $(U_\alpha, \varphi_\alpha) \in \mathcal{U}$ is itself an element of $\mathcal{U}$.

A C^∞ manifold is a topological manifold endowed with a C^∞ structure.

Definition 4.

Let M^d and N^n be two C^∞ manifolds. We say that $F : M \longrightarrow N$ is a C^∞ map if, for every $p \in M$, there exist coordinate neighborhoods (U, φ) of p and (V, ψ) of $F(p)$ with $F(U) \subset V$ such that the mapping $\psi \circ F \circ \varphi^{-1} : \varphi(U) \longrightarrow \mathbb{R}^n$ is C^∞.

We will write $C^\infty(p)$ for the algebra of C^∞ functions from some manifold M into $\mathbb{R}$ whose domain of definition includes some open neighborhood of $p \in M$, with the functions being identified if they agree on some neighborhood of p.

Definition 5.

We define the tangent space $T_p(M)$ to a manifold M at the point $p \in M$ to be the set of all linear mappings $X_p : C^\infty(p) \longrightarrow \mathbb{R}$ satisfying Leibniz rule, i.e.,

$$X_p(fg) = (X_p f)\, g(p) + f(p)\, (X_p g).$$

The tangent bundle of M, $T(M)$, is defined to be $\bigcup \{T_p(M) ; p \in M\}$.

Remark 1.

For each $p \in M$, $T_p(M)$ is a vector space over $\mathbb{R}$. Moreover, $T(M)$ can be adjoined a C^∞ structure making it a manifold (see Lemma 7.6.1 in Boothby (1986)).

Definition 6.

Let $F : M^d \longrightarrow N^n$ be a C^∞ map of manifolds. We define the differential of F at $p \in M$ to be the vector space homomorphism $F_* : T_p(M) \longrightarrow T_{F(p)}(N)$ specified by

$$F_*(X_p) f = X_p(f \circ F) \qquad \text{for } f \in C^\infty(p).$$

Remark 2.

Since $f \circ F \in C^\infty(p)$, this definition gives $F_*(X_p)$ as a map from $C^\infty(p)$ into $\mathbb{R}$.

If we take F to be the mapping φ from the coordinate neighborhood (U, φ) of p, we have $\varphi : M^d \longrightarrow \mathbb{R}^d$ and $\varphi_* : T_p(M) \longrightarrow T_{\varphi(p)}(\mathbb{R}^d)$.

Now, the tangent space of $\mathbb{R}^d$ at the point a is a vector space with natural basis $\left\{\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_d}\right\}$. Therefore, φ is a diffeomorphism between the manifolds M and $\mathbb{R}^d$ so that $\varphi_*^{-1} : T_a(\mathbb{R}^d) \longrightarrow T_{\varphi^{-1}(a)}(M)$ can be used to define a basis $\left\{E_{1p}, \ldots, E_{dp}\right\}$ for $T_p(M)$, where $E_{ip} \equiv \varphi_*^{-1}\left[\frac{\partial}{\partial x_i}\right]$. Such bases (which depend on p through φ) are called coordinate frames.

Remark 3.

The above discussion shows that we have the following series of equalities:

$$\dim(T_p(M)) = \dim(T_{\varphi(p)}(\mathbb{R}^d)) = \dim(\mathbb{R}^d) = \dim(M) = d.$$

In fact, φ_* is an isomorphism from $T_p(M)$ onto $T_{\varphi(p)}(\mathbb{R}^d)$.

Definition 7.

A vector field of class C^r on M is a function assigning to each point $p \in M$ a vector $X_p \in T_p(M)$ whose components in the frame of any local coordinate neighborhood (U, φ) are functions of class C^r on the domain U. In other words, a C^∞ vector field is a mapping of the form $X : M \longrightarrow T(M)$ with $X_p = \sum_{i=1}^{d} \alpha_i(p) E_{ip}$ where $\alpha_i : M \longrightarrow \mathbb{R}$ $(i = 1, \ldots, d)$ is C^∞ and $E_{ip} = \varphi_*^{-1}\left[\frac{\partial}{\partial x_i}\right]$. The term vector field will be used for C^∞ vector fields.

Using the above concepts, time homogeneous systems of differential equations on $\mathbb{R}^d$,

$$\dot{F}_i(t) = f_i(F(t)) \qquad (i = 1, \ldots, d),$$

can be written in the form

$$F_*\left[\frac{d}{dt}\right] = X_{F(t)},$$

where X is a vector field.

To see this, write $X_{F(t)} = \sum_{i=1}^{d} \alpha_i(F(t)) E_{iF(t)} = \sum_{i=1}^{d} \alpha_i(F(t)) \frac{\partial}{\partial x_i}$, with the last equality holding because the function F maps $\mathbb{R}$ to $\mathbb{R}^d$ and, as a manifold, $\mathbb{R}^d$ can always be assigned a single coordinate neighborhood, namely $(\mathbb{R}^d, Id)$. Hence, we get $E_{iF(t)} = Id_*^{-1}\left[\frac{\partial}{\partial x_i}\right] = \frac{\partial}{\partial x_i}$ (which does not depend on F(t)).

Identifying $\begin{bmatrix} f_1(t, F(t)) \\ \vdots \\ f_d(t, F(t)) \end{bmatrix}$ with $\sum_{i=1}^{d} f_i(F(t))\, e_i$, $\{e_i ; i = 1, \ldots, d\}$ a basis for $\mathbb{R}^d$, and further identifying $\{e_i ; i = 1, \ldots, d\}$ and $\left\{\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_d}\right\}$, we easily obtain $f_i(F(t)) = \alpha_i(F(t))$.

Similarly, $F_* : T_t(\mathbb{R}) \longrightarrow T_{F(t)}(\mathbb{R}^d)$ and, reasoning as above, $\left[\frac{d}{dt}\right]$ is a basis for $T_t(\mathbb{R})$ so that $F_* \left[\frac{d}{dt}\right]$ makes sense. Consequently, using Theorem 4.1.6, p. 110 in Boothby (1986), we get that the expression for $F_* \left[\frac{d}{dt}\right]$ in local (and here even global) coordinates is $\sum_{i=1}^{d} \dot{F}_i(t) \frac{\partial}{\partial x_i}$, which, by equating with the right hand side, gives the expression for the differential equation in $\mathbb{R}^d$.

Definition 8.

A field ϕ of C^∞ bilinear forms on a manifold M consists of a function assigning to each point $p \in M$ a bilinear form ϕ_p on $T_p(M)$, that is a map, linear in each variable separately, $\phi_p : T_p(M) \times T_p(M) \longrightarrow \mathbb{R}$, such that, for any coordinate neighborhood (U, φ), the functions $\alpha_{ij}(p) = \phi(E_{ip}, E_{jp})$ defined by ϕ and the coordinate frames $\left\{E_{1p}, \ldots, E_{dp}\right\}$ are of class C^∞.

Definition 9.

A manifold M on which there is defined a field of C^∞ bilinear form ϕ satisfying:

1) $\phi_p(X_p, Y_p) = \phi_p(Y_p, X_p)$ and (symmetric)

2) $\phi_p(X_p, Y_p) \geq 0$ (with equality if and only if $X_p = 0$) (positive definite)

is called a Riemannian manifold and ϕ is referred to as the Riemannian metric.

Remark 4.

Given any C^∞ manifold, it is always possible to define a C^∞ Riemannian metric (Theorem 5.4.5 in Boothby (1986)). Then, from this Riemannian metric, one can define a norm $\|X_p\| \equiv [\phi_p (X_p, X_p)]^{\frac{1}{2}}$ in each $T_p (M)$.

Now, any connected manifold is arc connected, i.e., for all a, $b \in M$, there exists a continuous function $f : [0, 1] \longrightarrow M$ such that $f(0) = a$ and $f(1) = b$. Moreover, f can be taken to be piecewise of class C^1 (see Buck (1978, p. 35 ; the proof there is given for $M = \mathbb{R}^d$ but the argument extends trivially)). Without loss of generality, we may assume that f itself is of class C^1 (otherwise, add up the pieces) and define the length of a curve on M to be

$$L = \int_0^1 \left[\phi_{p(t)} \left[\frac{dp}{dt}, \frac{dp}{dt}\right]\right]^{\frac{1}{2}} dt,$$

where $\frac{dp}{dt} \in T_{p(t)} (M)$ denotes the tangent vector to the curve at p (t), i.e., $\frac{dp}{dt} \equiv p_* \left[\frac{d}{dt}\right]$, $t \in (0, 1)$.

Theorem 1.

A connected Riemannian manifold is a metric space with the metric $\rho (p, q)$ defined to be the infimum of the length of the piecewise C^1 curves from p to q.
Its metric space topology and manifold topology agree.

Proof.

See Boothby (1986, Theorem 5.3.1, p. 189).

■

Remark 5.

If $M = \mathbb{R}^d$ and $T_p(M)$ is identified with $\mathbb{R}^d$, $\phi_p(x, y)$ can be defined to be $x'B_p y$, where B_p is a $d \times d$, symmetric, positive definite matrix, depending on p in a C^∞ way.

Now take $B_p = B$ for all $p \in M$ and recall that any positive definite (and in particular symmetric) matrix is in fact the variance–covariance matrix of d random variables $X_1, \ldots, X_d$ which can be taken to be multivariate normal. Therefore, we can regard any such multivariate normal distribution $N_d(0, \Sigma)$ as generating a metric on $\mathbb{R}^d$ with constant bilinear form Σ.

When one uses $\Sigma = \sigma^2 \mathrm{Id}$ (mimimization problems by ordinary least squares methods), this metric turns out to be the usual Euclidean metric (see the next example). Other variance–covariance matrices (generalized least squares) then simply correspond to other metrics on $\mathbb{R}^d$.

We now provide some examples illustrating the structure of some very common manifolds:

Example 1. The C^∞ manifold $\mathbb{R}^d$.

Using the Euclidean metric, $\mathbb{R}^d$ becomes a Hausdorff space with a countable basis of open sets and is trivially locally Euclidean. We can cover $\mathbb{R}^d$ with a single coordinate neighborhood $(\mathbb{R}^d, \mathrm{Id})$ which can be shown to define a C^∞ stucture on this space. The tangent space at $a \in \mathbb{R}^d$, $T_a(\mathbb{R}^d)$, is then a vector space whose basis $\{E_{1a}, \ldots, E_{da}\}$ is defined by $E_{ia} = \mathrm{Id}_*^{-1} \frac{\partial}{\partial x_i} = \frac{\partial}{\partial x_i}$. Since, in this case, we have only one coordinate neighborhood, these coordinate frames do not depend on a and we have $E_{ia} = E_i$ for all $a \in \mathbb{R}^d$. Moreover, $T_a(\mathbb{R}^d)$ can be identified with $\mathbb{R}^d$.

The field ϕ of C^∞ symmetric, positive definite bilinear form can be taken to be the natural inner product on $\mathbb{R}^d$, i.e., $\phi_a\left[\sum_{i=1}^{d} \alpha_i \frac{\partial}{\partial x_i}, \sum_{i=1}^{d} \beta_i \frac{\partial}{\partial x_i}\right] = \sum_{i=1}^{d} \alpha_i \beta_i$. Note that $\phi_a\left[\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\right] = \delta_{ij}$ (Kronecker δ).

The length of a curve in $\mathbb{R}^d$, say $p(t) = (x_1(t), \ldots, x_d(t))$, $0 \le t \le 1$, is then the usual formula for arc length:

$$L = \int_0^1 \left[\phi_{p(t)}\left[\frac{dp}{dt}, \frac{dp}{dt}\right]\right]^{\frac{1}{2}} dt = \int_0^1 \left[\phi_{p(t)}\left[\sum_{i=1}^{d} \frac{\partial x_i(t)}{\partial t} \frac{\partial}{\partial x_i}, \sum_{i=1}^{d} \frac{\partial x_i(t)}{\partial t} \frac{\partial}{\partial x_i}\right]\right]^{\frac{1}{2}} dt$$

$$= \int_0^1 \left[\sum_{i=1}^{d} \left[\frac{\partial x_i(t)}{\partial t}\right]^2\right]^{\frac{1}{2}} dt$$

Example 2. The C^∞ manifold of $d \times d$ real valued matrices, $gl(d, \mathbb{R})$.
This example follows from the previous one by simply identifying $gl(d, \mathbb{R})$ with $\mathbb{R}^{d^2}$.

Example 3. The C^∞ manifold $\mathbb{S}^{d-1} \equiv \{x \in \mathbb{R}_0^d : |x| = 1\}$ $(d \ge 2)$.
Taking $\mathbb{S}^{d-1}$ with the topology induced by $\mathbb{R}_0^d$ gives a Hausdorff space with a countable basis of open sets. There are several ways to show that $\mathbb{S}^{d-1}$ is locally Euclidean. For example, define, for $i = 1, \ldots, d$,

$$U_i^+ = \{(x_1, \ldots, x_d) \in \mathbb{S}^{d-1} : x_i > 0\}, \text{ open in } \mathbb{S}^{d-1}, \text{ and}$$
$$U_i^- = \{(x_1, \ldots, x_d) \in \mathbb{S}^{d-1} : x_i < 0\}, \text{ open in } \mathbb{S}^{d-1}.$$

Then $\mathbb{S}^{d-1} = \left[\bigcup_{i=1}^{d} U_i^+\right] \bigcup \left[\bigcup_{i=1}^{d} U_i^-\right]$ and the homeomorphisms $\rho_i^{\pm} : U_i^{\pm} \longrightarrow \mathbb{R}^{d-1}$, which are defined by

$$\rho_1^{\pm}(x_1, \ldots, x_d) = (x_2, \ldots, x_d),$$

$$\rho_2^{\pm}(x_1, \ldots, x_d) = (x_1, x_3, \ldots, x_d),$$

$$\vdots$$

$$\rho_d^{\pm}(x_1, \ldots, x_d) = (x_1, \ldots, x_{d-1}),$$

enable us to set up coordinate neighborhoods $(U_i^{\pm}, \rho_i^{\pm})$ which are C^{∞} compatible and define a C^{∞} structure on $\mathbb{S}^{d-1}$. Note that, in this case, the coordinate frames $\{E_{1s}, \ldots, E_{d-1s}\}$ on the tangent space at $s \in U_i^{\pm} \subset \mathbb{S}^{d-1}$, $T_s(\mathbb{S}^{d-1})$, will depend on s through $\rho_{i*}^{\pm}$.

By Corollary 5.2.5 in Boothby (1986, p. 185), the structure of $\mathbb{S}^{d-1}$ as a subset of $\mathbb{R}^d$ allows us to induce a Riemannian metric (i.e., a field of bilinear forms) on $\mathbb{S}^{d-1}$ from $\mathbb{R}^d$. When this Riemannian metric on $\mathbb{R}^d$ is taken to be the natural inner product (so that the metric on $\mathbb{R}^d$ is simply the Euclidean metric), $\mathbb{S}^{d-1}$ becomes a metric space with the usual metric defined by the arc length between two points. Indeed, in view of Theorem 1, this arc is the shortest piecewise C^1 curve between two points (i.e., the geodesic). A formal proof of this fact is nevertheless a little bit lengthy since, in this case, we have to work locally.

Example 4. The C^{∞} manifold of invertible matrices in $gl(d, \mathbb{R})$, $Gl(d, \mathbb{R})$.

As a general rule, any open subset of a C^{∞} manifold is itself a C^{∞} manifold (see Example 3.1.6, p. 56 in Boothby (1986)). This result provides a way of showing that $Gl(d, \mathbb{R})$ is a C^{∞} manifold. Indeed, $Gl(d, \mathbb{R})$ is the complement in the C^{∞} manifold

gl $(d, \mathbb{R})$ of the set $\mathcal{A} = \{A \in gl(d, \mathbb{R}) : \det(A) = 0\}$. Since det is a continuous map, $\mathcal{A} = (\det)^{-1}\{0\}$ is closed, which implies that Gl $(d, \mathbb{R})$ is open. Note that Gl $(d, \mathbb{R})$ is not a connected manifold but, in this work, we will focus on Gl $(d, \mathbb{R})$ as a (Lie) group acting on some other manifold M^d rather than as a state space.

2.4. Groups, Semigroups, and their Orbits

The notion of a Lie group acting on a manifold M as well as the notions of semigroups and orbits associated with such an action pervade in most parts of this work. This section is therefore devoted to a basic review of the related theory. A more detailed treatment of sone of these topics can be found in Boothby (1986).

Definition 1.

Let G be a group which is at the same time a differentiable manifold. Then G is a Lie group provided that the mapping $G \times G \longrightarrow G$ defined by $(x, y) \longmapsto x\,y$ and the mapping $G \longrightarrow G$ defined by $x \longmapsto x^{-1}$ are both C^∞ mappings.

Example 1.

The set Gl $(d, \mathbb{R})$ of nonsingular, $d \times d$, real valued matrices is a group with respect to matrix multiplication. Identified with an open subset of $\mathbb{R}^{d^2}$, it is clearly a differentiable manifold. Since the maps $(A, B) \longmapsto AB$ and $A \longmapsto A^{-1}$ are C^∞, Gl $(d, \mathbb{R})$ is a Lie group.

As a special case, $\mathbb{R}_0 = Gl(1, \mathbb{R})$ is a Lie group with respect to multiplication (by a nonzero number).

Since all Lie groups are topological groups, like them they will act on spaces, in particular on manifolds. Therefore, we can still talk about continuous and transitive action, as well as about the isotropy group at a point. Being a Lie group instead of simply a topological group will just entail the possibility of stronger results (for example, see Theorem 2.2.1 (d)).

Let $\theta : \mathbb{R} \times M \longrightarrow M$ be a C^∞ mapping satisfying:

1) $\theta_0 (x) = x$ for all $x \in M$ and

2) $\theta_t \circ \theta_s (x) = \theta_{t+s} (x) = \theta_s \circ \theta_t (x)$ for all $x \in M$.

In other words, θ represents the action of the Lie group $\mathbb{R}$ on the manifold M (see Definition 2.2.2) and is called a (global) one–parameter group action.

Definition 2.

The infinitesimal operator of θ is defined to be the C^∞ vector field X on M given, for each $x \in M$ and $f \in C^\infty (x)$, by

$$X_x (f) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} [f (\theta_{\Delta t} (x)) - f (x)].$$

Proposition 1.

Let θ be a C^∞ map as above and let X be the infinitesimal generator of θ. For fixed $x \in M$, write $F (t)$ for $\theta_t (x)$. Then:

a) X is invariant under the action of θ on M, i.e., $\theta_{t*} (X_x) = X_{\theta_t(x)}$ for all $t \in \mathbb{R}$, and

b) $F_* \left[\frac{d}{dt}\right]\Big|_{t = t_0} = X_{\theta_{t_0}(x)} = X_{F(t_0)}$.

Proof.

See Boothby (1986, Theorems 4.3.4 and 4.3.6). ■

Remark 1.

The result of this proposition should be compared to the discussion given in Subsection 2.3 on time homogeneous linear differential equations on $\mathbb{R}^d$. The point x is in fact the initial value for the solution of the differential equation.

The orbit of x under the group action of θ is defined to be the set

$$\{y \in M^d : y = \theta_t(x) = F(t) \text{ for some } t \in \mathbb{R}\}.$$

The curve $t \longmapsto F(t)$ defined on some open interval J of $\mathbb{R}$ is said to be an integral curve of X if $F_* \left[\frac{d}{dt}\right] = X_{\theta_t(x)} = X_{F(t)}$ on J, i.e., if this curve solves this differential equation on J .

So far we have shown that to a one–parameter group action θ on M corresponds a vector field on M called the infinitesimal generator of θ. A legitimate question would be to ask if this correspondence is unique, i.e., if to each vector field in T (M) corresponds a unique group action and conversely.

The answer to this question is "almost" yes. In general, with the above global definition of one–parameter group action, one cannot obtain such a one–to–one correspondence (see Boothby (1986, Example 4.3.9)). Nevertheless, if, loosely speaking, the condition $\theta_t \circ \theta_s(x) = \theta_{t+s}(x) = \theta_s \circ \theta_t(x)$ for all $x \in M$ is restricted to s and t belonging to open intervals depending on x, $(\alpha(x), \beta(x))$, then such a correspondence is true.

Group actions defined in such terms are called local one–parameter group actions or flows. A more complete discussion of such local actions is beyond our needs and the scope of this short introduction. A more detailed treatment can be found in Boothby (1986: see Definition 4.3.11 and Theorem 4.4.6).

The above results show that to any differential equation of the form $F_* \left[\frac{d}{dt}\right] = X_{\theta_t(x)} = X_{F(t)}$ corresponds a unique local one–parameter group action, which in turn determines the orbit of any initial value for this equation.

For example, when one looks at the linear time homogeneous differential equation on $\mathbb{R}^d$,

$$\dot{x}(t) = A\,x(t), \qquad A \in gl(d, \mathbb{R}),$$

the one–parameter group action (which is global in this case) is $\theta_t(x_o) = e^{At} x_o$ and the collection of mappings $\mathcal{G} = \{\theta_t(.) ; t \in \mathbb{R}\} = \{e^{At} ; t \in \mathbb{R}\}$ forms a group characterizing the integral curves of this differential equation.

Moreover, the orbit of some initial point x is obtained by applying the elements of $\mathcal{G}$ to x, i.e., if we denote this orbit by Gx, we have

$$Gx = \{y \in \mathbb{R}^d : y = e^{At} x \text{ for some } t \in \mathbb{R}\}.$$

In many cases, one wishes to restrict this situation to equations representing systems which evolve only forward (or backward) in time. This leads to the notion of semigroup which is defined as follows:

Definition 3.

Let E be a nonempty set endowed with a binary operation +. A subset E' of E is a semigroup if, for all $x, y, z \in E$,

1) $x + y \in E$ and

2) $x + (y + z) = (x + y) + z$.

If E" is an arbitrary subset of E, the group generated by E" is the smallest group in E which contains E". Similarly, the semigroup generated by E" is the smallest semigroup in E which contains E".

With this definition, one sees that the set $\mathcal{S} = \{e^{At} ; t \in \mathbb{R}_0^+\}$, obtained by restricting t to $\mathbb{R}^+$, is a semigroup which describes the behavior of the equation

$$\dot{x}(t) = A\,x(t)$$

forward in time. The positive orbit of some initial point $x \in \mathbb{R}^d$, which we denote by Sx, is then given by

$$Sx = \{y \in \mathbb{R}^d : y = e^{At}\,x \text{ for some } t \in \mathbb{R}_0^+\}.$$

Clearly, $\mathcal{G}$ from above is the group generated by this semigroup $\mathcal{S}$.

Using one–parameter group actions associated with the Lie group $\mathcal{U}$, i.e., in the discrete case of difference equations, the notion that the (semi)group associated with the equation should, when applied to some initial point, describe the (positive) orbit of this point immediately yields that, for the difference equation on $\mathbb{R}_0^d$

$$x_{n+1} = A\, x_n, \quad A \in Gl\,(d, \mathbb{R}),$$

the associated group and semigroup are

$$\mathcal{G} = \left\{ \prod_{k=0}^{n} A^{\sigma_k} ;\ \sigma_k = \pm 1,\ n \in \mathbb{N} \right\} \subset Gl\,(d, \mathbb{R}) \text{ and}$$

$$\mathcal{S} = \{A^n ;\ n \in \mathbb{N}_0\} \subset Gl\,(d, \mathbb{R}).$$

The ideas expressed above are easily extended to more complex situations and, in particular, to dynamical control systems which we now define.

Definition 4.

A continuous dynamical control system Σ is a 4-tuple $\Sigma = (M, U, \mathcal{U}, \mathcal{X})$, where

M^d is a manifold, constituting the state space,

$U \subset \mathbb{R}^n$ is a set of control values,

$\mathcal{U} \equiv \{u : \mathbb{R} \longrightarrow U\}$ is a set of admissible control functions, and

$\mathcal{X} \equiv \{X\,(u_t) \in T\,(M) ;\ u_t \in U\}$ is a collection of vector fields on M describing the dynamics of the system $x_* \left[\frac{d}{dt}\right] = X_{x_t}\,(u_t)$.

Discrete dynamical control systems are defined in the same way except that the control functions now map $\mathbb{N}$ into U and the dynamics is described by a difference equation $x_{n+1} = h\,(x_n, u_n)$ with $\mathcal{X} \equiv \{h\,(., u_n) \in Diff\,(M) ;\ u_n \in U\}$ is a collection of diffeomorphisms on M.

In the continuous time case on $\mathbb{R}^d$ and if the admissible control functions are piecewise constant, the group $\mathcal{G}$ and the semigroup $\mathcal{S}$ associated with a dynamical control system of the form $\dot{x}_t = A(u_t)\, x_t$, $A : U \longrightarrow gl(d, \mathbb{R})$, are then naturally defined by:

$$\mathcal{G} = \{e^{A(u_k)\, t_k} \ldots e^{A(u_1)\, t_1},\ t_i \in \mathbb{R} \text{ and } u_i \in U \text{ for } 1 \leq i \leq k \in \mathbb{N}\} \subset Gl(d, \mathbb{R}) \text{ and}$$

$$\mathcal{S} = \{e^{A(u_k)\, t_k} \ldots e^{A(u_1)\, t_1},\ t_i \in \mathbb{R}_0^+ \text{ and } u_i \in U \text{ for } 1 \leq i \leq k \in \mathbb{N}\} \subset Gl(d, \mathbb{R}).$$

For the more general situation given in the above definition of a continuous dynamical control system, the associated group and semigroup will be given by

$$\mathcal{G} = \{\exp(t_k X(u_k)) \ldots \exp(t_1 X(u_1)),\ t_i \in \mathbb{R} \text{ and } u_i \in U \text{ for } 1 \leq i \leq k \in \mathbb{N}\} \text{ and}$$

$$\mathcal{S} = \{\exp(t_k X(u_k)) \ldots \exp(t_1 X(u_1)),\ t_i \in \mathbb{R}_0^+ \text{ and } u_i \in U \text{ for } 1 \leq i \leq k \in \mathbb{N}\},$$

where $\exp(t\, X(u))$ can be regarded as a conventional notation for $\theta_u(t, .)$, with θ_u being the (local) one–parameter group associated with the vector field $X(u)$, i.e., $\exp(t_k X(u_k)) \ldots \exp(t_1 X(u_1))$ is defined recursively by

$$\exp(t_1 X(u_1)) = \theta_{u_1}(t_1, .) \text{ and}$$

$$\exp(t_k X(u_k)) \ldots \exp(t_1 X(u_1)) = \theta_{u_k}(t_k, \theta_{u_{k-1}}(t_{k-1}, .)).$$

Nevertheless, if M happens to be a Lie group, then the exp notation is more than a conventional notation. This will be explained at the end of the next subsection.

For a discrete dynamical control system on $\mathbb{R}^d_0$ of the form $x_{n+1} = A(u_n)\, x_n$, $A : U \longrightarrow Gl(d, \mathbb{R})$, the system's group and semigroup are in an analogous way given by

$$\mathcal{G} = \left\{ \prod_{k=0}^{n} A^{\sigma_k}(u_k) \text{ for } n \in \mathbb{N},\ u_k \in U,\ \text{and } \sigma_k = \pm 1 \right\} \subset Gl(d,\mathbb{R}) \text{ and}$$

$$\mathcal{S} = \left\{ \prod_{k=0}^{n} A(u_k) \text{ for } n \in \mathbb{N} \text{ and } u_k \in U \right\} \subset Gl(d,\mathbb{R}).$$

Finally, for a more general discrete dynamical control system, the associated group and semigroup will be given by

$$\mathcal{G} = \{ h^{\sigma_k \cdots \sigma_0}_{u_k \cdots u_0} \in \text{Diff}(M)\ ;\ \sigma_k = \pm 1 \text{ and } u_k \in U,\ 0 \le k \le n \in \mathbb{N} \},\ \text{and}$$

$$\mathcal{S} = \{ h_{u_k \cdots u_0} \in \text{Diff}(M)\ ;\ u_k \in U,\ 0 \le k \le n \in \mathbb{N} \},$$

where $h^{\sigma_k \cdots \sigma_0}_{u_k \cdots u_0}$ is defined by

$$h^{\sigma_i}_{u_i} = h(.,u_i),\ \text{for } \sigma_i = 1 \text{ or } h^{\sigma_i}_{u_i} = h^{-1}(.,u_i),\ \text{for } \sigma_i = -1,\ \text{and}$$

$$h^{\sigma_k \cdots \sigma_0}_{u_k \cdots u_0} = h^{\sigma_k}_{u_k} \circ h^{\sigma_{k-1}}_{u_{k-1}} \circ \cdots \circ h^{\sigma_0}_{u_0}.$$

Remark 2.

Note that, by the way semigroups were defined, the identity map may not be in $\mathcal{S}$ and hence it may be that $x \notin Sx$ (while, necessarily, $x \in Gx$).

2.5. Lie Groups and Lie Algebras

There is in fact a close correspondence between Lie groups and Lie algebras. Even though Lie algebras are not used in this work, for the sake of completeness and because Lie Algebras are widely used in the relevant literature, we will very briefly

talk about them and investigate their relationship to Lie groups. For more details, the reader should consult one of the classical references on this subject, e.g., Boothby (1986) and Helgason (1978), or, also, the nice review which can be found in MacDonald (1979).

If we denote by $V(M)$ the set of all C^∞ vector fields defined on the C^∞ manifold M, it is easy to see that $V(M)$ is itself a vector space over $\mathbb{R}$.

Definition 1.

A vector space L over $\mathbb{R}$ is a (real) Lie algebra if, in addition to its vector space structure, it possesses a product, that is a map $L \times L \longrightarrow L$ defined by $(X, Y) \longmapsto [X, Y] \in L$, with the following properties :

1) $[X, \alpha_1 Y_1 + \alpha_2 Y_2] = \alpha_1 [X, Y_1] + \alpha_2 [X, Y_2]$, (bilinearity)

2) $[X, Y] = -[Y, X]$, and (skew commutativity)

3) $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$. (Jacobi identity)

Example 1.

Taking $gl(d, \mathbb{R})$, the set of $d \times d$, real valued matrices, as a vector space over $\mathbb{R}$ with AB denoting the usual matrix multiplication of A and $B \in gl(d, \mathbb{R})$, it can easily be verified that the product $[A, B] = AB - BA$, called the commutator of A and B, will define a Lie algebra structure on $gl(d, \mathbb{R})$.

The same obviously applies to $\mathbb{R} = gl(1, \mathbb{R})$.

Example 2.

$V(M)$, defined as above, with the product $[X, Y] = XY - YX$ is also a Lie algebra (see Boothby (1986, Theorem 4.7.4, p. 153)).

Now we will investigate the relationship between Lie algebras and Lie groups mentioned at the beginning of this section.

Definition 2.

Let G be a Lie group. Then, for each $x \in G$, the left translation by a is the diffeomorphism $L_a : G \longrightarrow G$ defined by $L_a(x) = a\,x$.
If X is a vector field on G, we say that X is left invariant if X commutes with left translation, i.e., if

$$X \circ L_a = L_{a*} \circ X \text{ for all } a \in G.$$

Now, in the previous section, we have established that there exists a one–to–one correspondence between vector fields on G and flows $\theta_t : G \longrightarrow G$. Then for each $x \in G$, the map $\theta_{t*} : T_x(G) \longrightarrow T_{\theta_t(x)}(G)$ will allow us to compare, for each x, the rate of change of another vector field Y on G along the integral curve of X passing through x. This rate of change is denoted by $L_X Y$ and is formally defined hereafter.

Definition 3.

The C^∞ vector field $L_X Y$, called the Lie derivative of Y with respect to X, is defined, at each $x \in G$, by the limit

$$(L_X Y)_x = \lim_{t \to 0} \frac{1}{t} \left[Y_x - \theta_{t*} Y_{\theta(-t, x)} \right].$$

Theorem 1.

If G is a Lie group, then the left invariant vector fields on G form a Lie algebra $\mathcal{g}$ with Lie bracket $[X, Y] = L_X Y$ and dim (G) = dim ($\mathcal{g}$) (dim (G) as a manifold and dim ($\mathcal{g}$) as a vector space over $\mathbb{R}$). Moreover, if $F : G_1 \to G_2$ is a homomorphism of Lie groups, $F_* : \mathcal{g}_1 \to \mathcal{g}_2$ is a homomorphism of Lie algebras.

Proof.

See Boothby (1986, Corollary 4.7.10) and MacDonald (1979, pp. 97–98). ■

Remark 1.

Each $X \in \mathcal{g}$ is determined by its value $X_e \in T_e (G)$ at the identity element e of G because

$$X_x = (X \circ L_x)_e = \left[L_{x*} X \right]_e.$$

Conversely, each tangent vector $X \in T_e (G)$ determines a left-invariant vector field Y on G by the rule

$$Y_x = \left[L_{x*} X \right]_e.$$

On the other hand, when we have $\mathcal{g}$, the Lie algebra of the Lie group G, it is often possible to recover G from $\mathcal{g}$. This is done through the exponential mapping applied to the vector fields in $\mathcal{g}$. The basic idea is as follows.

Definition 4.

A one–parameter subgroup F of a Lie group G is a C^∞ homomorphism from $\mathbb{R}$ to G ($\mathbb{R}$ with the additive and G with the multiplicative structure).

Let F be a one–parameter subgroup of the Lie group G. Then

$$F_* : T_0(\mathbb{R}) \longrightarrow T_{F(0)}(G) = T_e(G),$$

i.e., upon identification of $T_0(\mathbb{R})$ with $\mathbb{R}$ and of $T_e(G)$ with $\mathcal{g}$, we have F_* as a homomorphism from $\mathbb{R}$ to $\mathcal{g}$. Moreover, given $X_e \in \mathcal{g} \cong T_e(G)$, the equation

$$F_* \left[\frac{d}{dt}\right]\Big|_0 = X_e$$

has a unique C^∞ homomorphism solution which we will denote by $F_{X_e}(t)$ (see Helgason (1978, Corollary 2.1.5)).

Definition 5.

For each $X \in \mathcal{g}$, we define $\exp(tX) = F_X(t)$, where $F_X : \mathbb{R} \longrightarrow G$ is the unique C^∞ homomorphism solving

$$F_* \left[\frac{d}{dt}\right]\Big|_0 = X_e.$$

In particular, $\exp X = F_X(1)$.

So, for $X \in \mathcal{g}$ and by the Inverse Function Theorem, exp X is a C^∞ homomorphism from an open neighborhood of $0 \in \mathcal{g}$ to an open neighborhood of e, the identity element in G. Nevertheless, it is in general not true that $G = \exp(\mathcal{g})$, i.e., the exponential map is usually not surjective. If G is connected, then $\mathcal{g}$ will give a chart of G around e and hence will generate G. If G is compact or Abelian (and

connected), then exp is a surjective map.

Theorem 2.

Consider the differential equation $F_* \left[\frac{d}{dt}\right] = X_{F(t)}$ on the Lie group G where X is a left invariant vector field.

Then the one–parameter group action of $\mathbb{R}$ on G which has the infinitesimal generator $X \in \mathcal{g}$, $\theta : \mathbb{R} \times G \longrightarrow G$, is given by $\theta(t, x) = R_{F(t)}(x)$, where the map $F : \mathbb{R} \longrightarrow G$ is the one–parameter subgroup of G given by $F(t) = \theta(t, e) = \exp(t\, X_e)$ and R_a is the right translation by a on G.

Proof.

See Boothby (1986, Theorems 4.5.13 and 4.6.9). ■

Hence, if the manifold M is also a Lie group, the exponential notation we used for the group $\mathcal{G} = \{\exp(t_k X(u_k)) \ldots \exp(t_1 X(u_1)),\ 1 \leq i \leq k \in \mathbb{N},\ t_i \in \mathbb{R},\ u_i \in U\}$ in Subsection 2.4 becomes more than a convention provided that $\exp(t_i X(u_i))$ is understood to mean $\exp(t_i X_e(u_i))$, with e denoting the identity elememt of M. Indeed, $\exp(t_i X(u_i))$ then corresponds directly to the one–parameter subgroup of M induced by the the vector field $X(u_i)$.

The idea is similar (if M is a Lie group and all vector fields are left invariant) for the expression $\exp(t_k X(u_k)) \ldots \exp(t_1 X(u_1))$ arising from our dynamical control system, except for the fact that the free choice of the u_i's and of the length of time t_i the vector field $X(u_i)$ is applied renders this expression simply a subgroup of M and not a one–parameter subgroup. Also note that, the dynamical control systems arising in this thesis (from stochastic linear difference or differential

equations) do not give left invariant vector fields.

We conclude this section by an example of a Lie group and its associated Lie algebra.

Example 3.

Consider the Lie group Gl $(d, \mathbb{R})$ of invertible real–valued $d \times d$ matrices. Then $\mathcal{gl}(d, \mathbb{R})$, the Lie algebra associated with Gl $(d, \mathbb{R})$ can be identified with gl $(d, \mathbb{R})$, the set of all real–valued $d \times d$ matrices.

Moreover, the mapping exp : $\mathcal{gl}(d, \mathbb{R}) \longrightarrow$ Gl $(d, \mathbb{R})$ is nothing but the usual exponential map associated with matrices, i.e., $e^A = \sum_{n=0}^{\infty} A^n / n!$.

Note however that $\det(e^A) = e^{\text{trace}(A)} > 0$, and hence that this exponential map is not surjective since its range is limited to the matrices in Gl $(d, \mathbb{R})$ with positive determinant.

For a proof of these statements, see Helgason (1978, pp. 110–111).

2.6. Markov Processes

The purpose of this section is to review some notions pertaining to Markov processes. The following facts can be found in greater details in several standard books on the subject, e.g., Ethier and Kurtz (1986) and Revuz (1984).

Let E be some metric space and let $\mathcal{T}$ be the topology induced on E by its

metric. Let $\mathcal{B}(E)$ be the σ-algebra of Borel subsets of E, i.e., the σ-algebra generated by the elements of $\mathcal{T}$. Then $(E, \mathcal{B}(E))$ is a measurable space.

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let T denote a discrete or continuous time set. T is a subset of $\mathbb{R}$ and hence, with the same notation as above, $(T, \mathcal{B}(T))$ is a measurable space.

Definition 1.

A stochastic process $X \equiv \{X(t,.) ; t \in T\}$ defined on $(\Omega, \mathcal{F}, P)$ and taking values in $(E, \mathcal{B}(E))$ is said to be $(\mathcal{B}(T) \times \mathcal{F})$ – measurable if $\{(t, \omega) \in T \times \Omega : X(t, \omega) \in B\}$ belongs to $\mathcal{B}(T) \times \mathcal{F}$ for all $B \in \mathcal{B}(E)$.

A collection $\{\mathcal{F}_t ; t \in T\}$ of sub-σ-algebras of $\mathcal{F}$ is said to be a filtration if $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ for all $s \leq t$.

The process X is said to be $\{\mathcal{F}_t\}$ – adapted with respect to $\{\mathcal{F}_t\}$ if, for all fixed $t \in T$ and for all $B \in \mathcal{B}(E)$, the set $\{\omega \in \Omega : X(t, \omega) \in B\}$ belongs to $\mathcal{F}_t$.

Remark 1.

We will denote by $\sigma(X(t,.))$ the smallest sub-σ-algebra of $\mathcal{F}$ which makes $X(t,.)$ measurable. From now on, $\sigma(X(t,.))$ will be abbreviated $\sigma(X_t)$ and $X(t, \omega)$, X_t. If we denote by $\mathcal{F}_t^X \equiv \sigma(X(s,.) ; s \leq t)$ the smallest sub-σ-algebra of $\mathcal{F}$ which makes $X(s,.)$ measurable for all $s \leq t$, it is clear that X is $\{\mathcal{F}_t\}$ – adapted if and only if $\mathcal{F}_t^X \subset \mathcal{F}_t$ for all $t \in T$.

Also note that, if T is a discrete subset of $\mathbb{R}$, having X measurable is equivalent to having X $\{\mathcal{F}_t\}$ – measurable for all $t \in T$ (i.e., $\{\omega : X(t, \omega) \in B\} \in \mathcal{F}_t$ for all $B \in \mathcal{B}(E)$ and for all $t \in T$) because T is countable. In the continuous case, having X

measurable is a stronger statement.

Definition 2.

Let $X \equiv \{X_n ; n \in \mathbb{N}\}$ be a discrete process defined on $(\Omega, \mathcal{F}, P)$ with values in $(E, \mathcal{B}(E))$. Assume X is adapted to the filtration $\{\mathcal{F}_n ; n \in \mathbb{N}\}$. Then X is said to be a Markov chain with respect to $\{\mathcal{F}_n\}$ if

$$P(X_{m+n} \in B \mid \mathcal{F}_m) = P(X_{m+n} \in B \mid \sigma(X_m))$$

for all m, $n \geq 0$ and for all $B \in \mathcal{B}(E)$.

If $\mathcal{F}_n = \mathcal{F}_n^X$ for all n, we simply say that X is an E-valued Markov chain on $(\Omega, \mathcal{F}, P)$.

Let $\mathcal{M}(E)$ and $\mathcal{P}(E)$ denote the sets of all measures and of all probability measures on $(E, \mathcal{B}(E))$ respectively. Write $B(E)$ for the set of all bounded real valued measurable functions on E. Then $B(E)$ is a Banach space with the supremum norm.

Definition 3.

A function $\mu(.,.)$ defined on $E \times \mathcal{B}(E)$ is called a transition function if, for all $x \in E$,

$$\mu(x,.) \in \mathcal{P}(E),$$

and, for all $B \in \mathcal{B}(E)$,

$$\mu(.,B) \in B(E).$$

A transition function $\mu(.,.)$ is a transition function for a time homogeneous (discrete) Markov process X if, for all $B \in \mathcal{B}(E)$,

$$P\left[X_{n+1} \in B \mid \mathcal{F}_n^X\right] = \mu(X_n, B)$$

or, equivalently, if, for all $f \in B(E)$,

$$E\left[f(X_{n+1}) \mid \mathcal{F}_n^X\right] = \int_E f(x)\ \mu(X_n, dx).$$

Remark 2.

The first equation above means that, when the values taken by $\{X_0, \ldots, X_n\}$ are known, the probability that the random variable X_{n+1} (i.e., $X(n+1, \omega)$) will take a value in the set $B \in \mathcal{B}(E)$ is given by $\mu(x_n, B)$ where x_n denotes the known value of X_n. Clearly, if $X_m = x_m$, $P\left[X_{m+1} \in B \mid \mathcal{F}_m^X\right] = \mu(x_m, B)$ and therefore, the probability that, at the next step, it will end up in B depends only on its last position, and not on the time this transition is made. This justifies the term "time homogeneous".

Now $\mu(x, B)$ is really the one-step transition probability from x to B. For this reason, we will sometimes write $\mu(1, x, B)$ instead of $\mu(x, B)$. This will allow us to use the notation $\mu(n, x, B)$ for the n-step transition probability from x to B, i.e., for all $B \in \mathcal{B}(E)$,

$$P\left[X_{m+n} \in B \mid \mathcal{F}_m^X\right] = \mu(n, X_m, B).$$

Definition 4.

A transition function is said to satisfy the Chapman – Kolmogorov property if

$$\mu(n + m, x, B) = \int_E \mu(n, y, B)\ \mu(m, x, dy),$$

where

$\int_E \mu(n, y, B)\ \mu(m, x, dy)$ can be written as

$$\int_E \cdots \int_E \mu(y_{n-1}, B)\ \mu(y_{n-2}, dy_{n-1}) \cdots \mu(y_1, dy_2)\ \mu(m, x, dy_1).$$

Remark 3.

Note that the transition function of a homogeneous Markov chain (and of Markov processes in general) satisfies the Chapman – Kolmogorov property since

$$\begin{aligned}
\mu(n+m, X_u, B) &= P\left[X_{n+m+u} \in B \mid \mathcal{F}_u^x\right] \\
&= E\left[P\left[X_{n+m+u} \in B \mid \mathcal{F}_{m+u}^x\right] \,\middle|\, \mathcal{F}_u^x\right] \\
&= E\left[\mu(n, X_{m+u}, B) \mid \mathcal{F}_u^x\right] \\
&= \int_E \mu(n, y, B)\ \mu(m, X_u, dy).
\end{aligned}$$

Definition 5.

The probability measure $\nu \in \mathcal{P}(E)$ given by $\nu(B) = P(X_0 \in B)$ for $B \in \mathcal{B}(E)$ is called the initial distribution of the Markov process X.

Repeatedly using the Markov property, it can be shown that the transition function of a homogeneous Markov process X and its initial distribution ν uniquely determine the finite-dimensional distributions of X by

$$P(X_0 \in B_0, X_{n_1} \in B_1, \ldots, X_{n_k} \in B_k) = \qquad (*)$$

$$\int_{B_0} \cdots \int_{B_{k-1}} \mu(n_k - n_{k-1}, y_{k-1}, B_n)\ \mu(n_{k-1} - n_{k-2}, y_{k-2}, dy_{k-1}) \cdots \mu(n_1, y_0, dy_1)\ \nu(dy_0)$$

Conversely, we have the following theorem:

Theorem 1.

If E is a complete separable space and $\mu(.,.)$ is a transition function satisfying the Chapman – Kolmogorov property, then there is a Markov chain X defined on E whose finite–dimensional distributions are uniquely determined by (*) above.

Proof.

See Ethier and Kurtz (1986, Theorem 4.1.1, p. 157). ■

Remark 4.

Manifolds are Polish spaces, i.e., they are locally compact with countable basis (LCCB) (see Subsection 2.3). Hence they are separable and complete by Hewitt and Stromberg (1965, Theorems 2.6.22, p. 61 and 2.6.50, p. 67) (for completeness, the proof in M is similar to the proof in $\mathbb{R}^d$ given there).

In order to develop further concepts related to Markov chains, we need some additional definitions. In the following, let M (E) denote the set of all real valued $\mathcal{B}$– measurable functions on E.

Definition 6.

A kernel on $(E, \mathcal{B}(E))$ is a mapping K from $E \times \mathcal{B}(E)$ into $(-\infty, \infty]$ such that

1) $K(., B) : E \rightarrow (-\infty, \infty]$ is a $\mathcal{B}$– measurable function for all $B \in \mathcal{B}(E)$ and

2) $K(x, .) : \mathcal{B}(E) \longrightarrow (-\infty, \infty]$ is a measure on $(E, \mathcal{B}(E))$ for all $x \in E$.

The kernel K is said to be positive if its range is $[0, \infty]$. It is said to be σ-finite if all the measures $K(x, .)$ are σ-finite. It is said to be proper if E is the union of an increasing sequence $\{E_n ; n \geq 0\}$ of subsets of E such that the functions $K(., E_n)$ are bounded. It is said to be bounded if $|K(x, B)| \leq M < \infty$ for all $x \in E$ and all $B \in \mathcal{B}(E)$.

Remark 5.

A bounded kernel is a proper kernel and a proper kernel is σ-finite.

The kernel $K(., .)$ can be used to define an operator K from $M(E)$ to $M(E)$ by

$$K f(x) = \int_E f(y) K(x, dy).$$

Also, if $m \in \mathcal{P}(E)$ and $B \in \mathcal{B}(E)$, we define

$$m K(B) = \int_E K(x, B) m(dx).$$

Definition 7.

The measure $m \in \mathcal{P}(E)$ will be said to be invariant under the kernel K if, for all $B \in \mathcal{B}(E)$, we have

$$m(B) = m K(B).$$

Remark 6.

Transition probabilities are simply positive kernels satisfying the condition $K(x, E) = 1$ for all $x \in E$, i.e., $K(x, .)$ is a probability measure for all $x \in E$.

Moreover, if $\mu(.,.)$ is a one step transition probability for a Markov chain, we have $\mu(x, B) = \mu I_B(x)$, where $I_B(.)$ denotes the indicator function of the set $B \in \mathcal{B}(E)$.

Associated with Markov processes are also the notions of semigroup and generator, which we now define, with the time set $T = \mathbb{R}^+$ or $\mathbb{N}$.

Definition 8.

A one–parameter family $\{T(t) ; t \in T\}$ of bounded linear operators on a Banach space $\mathcal{E}$ is called a semigroup if

1) $T(0) = I$, the identity map, and

2) $T(s + t) = T(s)\, T(t)$ for all $s, t \in T$.

A semigroup $\{T(t)\}$ is said to be strongly continuous if $\lim_{t \to 0} T(t) f = f$ for every $f \in \mathcal{E}$ (the limit being taken in the Banach space norm).

It is said to be contracting if $\|T(t)\| \leq 1$ for all $t \in T$, where $\|.\|$ is the usual operator norm.

In the discrete case of Markov chains and by the Chapman – Kolmogorov property,

$$T(n) f(x) = \int_E f(y)\, \mu(n, x, dy)$$

defines a measurable (for each n) and contracting semigroup on the Banach space of bounded measurable functions (on the state space of the Markov chain).

Definition 9.

The X be a Markov chain with state space E and let $\{T(n) ; n \in \mathbb{N}\}$ be a discrete time semigroup on the Banach space $\mathcal{E} = B(E)$ of bounded measurable functions on E. Then the Markov chain X corresponds to $\{T(n)\}$ if, for all $f \in B(E)$,

$$E[f(X_{n+k}) | \mathcal{F}_n^X] = T(k) f(X_n).$$

Remark 7.

If $\{T(n)\}$ is defined by $T(n) f(x) = \int_E f(y) \mu(n, x, dy)$, the definition above is simply the definition of a Markov chain.

Theorem 2.

If X is a Markov chain with initial distribution ν and corresponding to the semigroup $\{T(n) ; n \in \mathbb{N}\}$, then the finite dimensional distributions of X are fully determined by ν and $\{T(n)\}$.

Proof.

See Ethier and Kurtz (1986, Proposition 4.1.6). ■

Definition 10.

The (infinitesimal) generator of a semigroup $\{T(t) ; t \in T\}$ on a Banach space $\mathcal{E}$ is the linear operator A on $\mathcal{E}$ defined by

$$A f = \lim_{t \to 0} \frac{1}{t} [T(t) f - f],$$

the limit being taken in the topology of the Banach space norm.

The domain $\mathcal{D}(A)$ of A is the subspace of all $f \in \mathcal{E}$ such that this limit exists.

Under some conditions which are given below, the Hille – Yosida theorem says that a strongly continuous contraction semigroup can be recovered from its generator. The result goes as follows:

Definition 11.

Let A be a linear operator on the Banach space $\mathcal{E}$ and assume that A is closed, i.e., that the graph of A is closed as a topological subspace of $\mathcal{E} \times \mathcal{E}$. If, for some real number λ, we have

1) the operator $\lambda I - A$ is one–to–one,

2) the range of $\lambda I - A$ is $\mathcal{E}$, and

3) $(\lambda I - A)^{-1}$ is a bounded linear operator on $\mathcal{E}$,

then λ is said to belong to the resolvent set $\rho(A)$ of A and $R_\lambda \equiv (\lambda I - A)^{-1}$ is called the resolvent (at λ) of A.

Definition 12.

A linear operator A on the Banach space $\mathcal{E}$ is said to be dissipative if

$$\|\lambda f - A f\| \geq \lambda \|f\| \text{ for every } f \in \mathcal{D}(A) \text{ and } \lambda > 0.$$

Theorem 3. (Hille – Yosida)

A linear operator A on the Banach space $\mathcal{E}$ is the generator of a strongly continuous contraction semigroup $\{T(t)\}$ on $\mathcal{E}$ if and only if

1) $\mathcal{D}(A)$ is dense in $\mathcal{E}$,

2) A is dissipative, and

3) the range of $\lambda I - A$ is $\mathscr{E}$ for some positive λ in the resolvent set $\rho(A)$.

When these conditions are satisfied, $\{T(t)\}$ is, for all $t \in T$ and all $f \in \mathscr{E}$, defined by the uniform limit on bounded intervals

$$T(t) f = \lim_{\lambda \to \infty} \exp(t A_\lambda) f,$$

where A_λ is called the Yosida approximation of A and defined by

$$A_\lambda = \lambda A R_\lambda = \lambda A (\lambda I - A)^{-1}.$$

Proof.

See Ethier and Kurtz (1986, Theorem 1.2.6 and Proposition 1.2.7). ■

2.7. Irreducibility and Recurrence of Markov Chains

Irreducibility and recurrence notions will become quite important later on. Therefore, this section is devoted to a brief review of these topics. In this, we follow mainly Jain and Jamison (1967), Stettner (1988), and Tweedie (1976).

Definition 1.

Let $(E, \mathscr{B})$ be a measurable space. We say that $\mathscr{B}$ is separable if $\mathscr{B}$ is generated by a countable subclass of subsets of E.

Remark 1.

If E is an L.C.C.B. space and $\mathcal{B}$ is the Borel σ-algebra of E, i.e., $\mathcal{B} = \mathcal{B}(E)$, then $\mathcal{B}$ is clearly separable.

From now on, we will always assume that our σ-algebra $\mathcal{B}$ is separable. In fact, in most cases, $\mathcal{B}$ will be the Borel σ-algebra $\mathcal{B}(E)$.

Next, we introduce the terminology and the notations used in the text. For each $B \in \mathcal{B}(E)$ and each $x \in E$, define

$$\tau(x, B) = \inf(n > 0 ; X_n \in B \mid X_0 = x),$$

$$F(n, x, B) = P(\tau(x, B) = n),$$

$$L(x, B) = P\left[\bigcup_{n=1}^{\infty} \left[X_n \in B\right] \;\middle|\; X_0 = x\right] = \sum_{n=1}^{\infty} F(n, x, B),$$

$$G(x, B) = \sum_{n=1}^{\infty} \mu(n, x, B), \text{ and}$$

$$Q(x, B) = P(X_n \in B \text{ i.o.} \mid X_0 = x)$$

$$= P\left[\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \left[X_k \in B\right] \;\middle|\; X_0 = x\right].$$

In other words, $\tau(x, B)$ denotes the first hitting time for the set B starting from x, $F(n, x, B)$ is the probability that the first entrance in B starting from x takes place at the n^{th} step, $L(x, B)$ is the probability that the chain ever reaches B starting

from x, and Q (x, B) denotes the probability of returning to B infinitely often starting from x.

Remark 2.

For all B $\in \mathcal{B}(E)$, we have:

1) L (x, B) > 0 if and only if G (x, B) > 0 (Jain and Jamison (1967, p. 3–4)), and
2) L (x, B) =1 for all x $\in E$ if and only if Q (x, B) = 1 for all x $\in E$ (Tweedie (1976, Proposition 3.1))

Definition 2.

A set B $\in \mathcal{B}(E)$ is called invariant (or stochastically closed) if B $\neq \phi$ and μ (x, B) = 1 for all x $\in$ B.

An invariant set which does not contain two disjoint invariant sets is called indecomposable.

Definition 3.

A set B $\in \mathcal{B}(E)$ such that Q (x, B) = 0 for all x $\in E$ is called inessential. Otherwise, it is called essential.

An essential set which is the union of countably many inessential sets is said to be improperly essential. Otherwise it is called properly essential.

Remark 3.

Inessential sets are those in which the process X will only stay for finitely many steps with probability one.

If B is inessential or improperly essential, then B is the countable union of strongly transient sets, i.e., $B = \bigcup_{n=1}^{\infty} B_n$ with, for all n and all $x \in E$,

$$G(x, B_n) < \infty.$$

Also, strongly transient sets are inessential (but the converse is false). For more details, see Jain and Jamison (1967, p. 34 as well as Corollaries 4.1 and 4.2).

Definition 4.

Let $B \in \mathcal{B}(E)$ be invariant and φ be a σ-finite measure on $(E, \mathcal{B}(E))$ with $\varphi(B) > 0$. Then B is called φ-irreducible if $L(x, F) > 0$ for all $x \in B$ whenever $F \subset B$ and $\varphi(F) > 0$.

A set $D \in \mathcal{B}(E)$ is called weakly recurrent if $G(x, D) = \infty$ for all $x \in E$. D is called strongly recurrent if $L(x, D) = Q(x, D) = 1$ for all $x \in E$.

Definition 5.

Let φ be a σ-finite nontrivial measure on $(E, \mathcal{B}(E))$. The Markov chain X is said to be φ-irreducible if, for all $x \in E$ and $B \in \mathcal{B}(E)$, $G(x, B) > 0$ (or equivalently, by Remark 2, if $L(x, B) > 0$) whenever $\varphi(B) > 0$.

The φ-irreducible chain X is said to be (weakly) recurrent if, for all $x \in E$ and $B \in \mathcal{B}(E)$, $G(x, B) = \infty$ whenever $\varphi(B) > 0$. The same chain X is φ-recurrent or strongly recurrent if, for all $x \in E$ and $B \in \mathcal{B}(E)$, $L(x, B) = 1$ whenever $\varphi(B) > 0$.

If X is φ-recurrent for some such φ, then X is said to be Harris recurrent.

Remark 4.

For all $B \in \mathcal{B}(E)$, $Q(x, B) = 1$ for all $x \in E$ (strong recurrence) implies $G(x, B) = \infty$

for all $x \in E$ (weak recurrence). The reverse implication in not true (Tweedie (1976, Proposition 3.2)).

Proposition 1.

Suppose X is φ-irreducible for some φ. Then there is a measure $\tilde{\varphi}$ on $\mathcal{B}(E)$ which is a maximal irreducibility measure for X in the sense that

1) X is $\tilde{\varphi}$-irreducible,

2) if X is ψ-irreducible for some ψ, then $\psi << \tilde{\varphi}$, and

3) for all $B \in \mathcal{B}(E)$, $\tilde{\varphi}(B) = 0$ implies $\tilde{\varphi}\{x \in E : L(x, B) > 0\} = 0$.

Proof.

See Tweedie (1976, Proposition 2.1). ■

Definition 6.

The Markov chain X is said to be strongly φ-irreducible if it is φ-irreducible and $\tilde{\varphi}$ as in Proposition 1 also satisfies

4) for all $B \in \mathcal{B}(E)$, $\tilde{\varphi}(B) = 0$ implies $\{x \in E : L(x, B) > 0\} = \phi$.

Definition 7.

An invariant set $B \in \mathcal{B}(E)$ is said to be φ-minimal for $\varphi \in \mathcal{M}(E)$ if it does not contain any invariant set of φ-measure less than $\varphi(B)$.

Write, for $B \in \mathcal{B}(E)$ and $x \in E$,

$$R(x, B) = \sum_{n=1}^{\infty} n F(n, x, B),$$

i.e., if $\tau(x, B)$ is a proper random variable (i.e., if $1 = P(\tau(x, B) < \infty) = L(x, B)$), then $R(x, B) = E(\tau(x, B))$.

Definition 8.

If, for $B \in \mathcal{B}(E)$ and for all $x \in E$, $L(x, B) = 1$ and $R(x, B) < \infty$, the set B is called positive.

If $\overline{\lim_{n\to\infty}} \mu(n, x, B) > 0$ for all $x \in E$, B is called weakly positive.

A set $B \in \mathcal{B}(E)$ is called null if $R(x, B) = \infty$ for all $x \in B^C$.

B is called strongly null if $\lim_{n\to\infty} \mu(n, x, B) = 0$ for all $x \in E$.

Remark 5.

$B \in \mathcal{B}(E)$ positive does not necessarily imply B weakly positive (Tweedie (1976, p. 747, and 1975, Section 6)).

If B is positive and $\sup \{R(x, B) ; x \in B\} \leq k$ for some $k < \infty$, then B is weakly positive (Tweedie (1976, Proposition 4.1)).

Definition 9.

If the Markov chain X is φ-irreducible, X is said to be positive if, for $B \in \mathcal{B}(E)$, $\varphi(B) > 0$ implies B is weakly positive (i.e., positive chains are weakly recurrent). Otherwise X is called null.

The Markov chain X is called ergodic if it is positive. (Also see Remark 2.8.3 (1).)

Remark 6.

If X is a (φ-irreducible) positive chain, then positivity of the chain X implies that X is weakly recurrent (with respect to φ) and hence (by Theorem 2.8.2 (c)) that X admits a unique invariant probability measure, λ. From Remark 2.8.2, it follows that X can be restricted to a φ-recurrent (strongly recurrent) chain by removing a λ-null set. Moreover, Theorem 2.8.2 (b) states that, for a positive chain and for every set $B \in \mathcal{B}(E)$ with $\varphi(B) > 0$, $R(x, B) < \infty$ λ-a.s. Hence, by removing a λ-null set from the state space of X, we get that all sets $B \in \mathcal{B}(E)$ with $\varphi(B) > 0$ are positive, i.e., $R(x, B) < \infty$ and $L(x, B) = 1$ for all $x \in E \setminus K$ with $\lambda(K) = 0$. In other words, the expected time for the first entry from $x \in E \setminus K$ in any set $B \in \mathcal{B}(E)$ with $\varphi(B) > 0$ is finite.

Definition 10.

A sequence of d disjoint sets $\{C_n ; 1 \le n \le d\}$ in $\mathcal{B}$ is called a cycle if, for $1 \le j \le d-1$,

$$\mu(x, C_{j+1}) = 1, \qquad x \in C_j, \text{ and}$$
$$\mu(x, C_1) = 1, \qquad x \in C_d.$$

Theorem 1.

Assume that $\mathcal{B}$ is separable and let X be a Markov chain valued in $(E, \mathcal{B})$.
If X is φ-irreducible, then there is a cycle $\{C_n ; 1 \le n \le d\}$ for which the following holds:

a) Let $K = \bigcup_{i=1}^{d} C_n$, $C_i \cap C_j = \phi$ for $i \neq j$. Then K^c is inessential or improperly essential and $\varphi(K^c) = 0$.

b) If $\{C'_m ; 1 \le m \le d'\}$ is a cycle, then d' is a divisor of d and, for each $1 \le m \le d'$,

C_m' differs from a union of d/d' members of $\{C_n ; 1 \leq n \leq d\}$ by a φ–null set which is either inessential or improperly essential.

Proof.

See Theorem 2.1 in Jain and Jamison (1967). ■

Theorem 2.

Suppose that $(E, \mathcal{B})$ is the measurable state space for the Markov chain X. Assume that $\mathcal{B}$ is separable and that X is φ-irreducible for some σ-finite measure φ.

Then either

- $E = H \cup I$, where $H \cap I = \phi$, H is invariant and φ-recurrent, and I is inessential or improperly essential with $\varphi(I) = 0$, or
- $E = \bigcup_{i=1}^{\infty} S_i$, where all the S_i's are strongly transient, i.e., $G(x, S_i) < \infty$ for all i and all $x \in E$.

Moreover, the first case holds only when $\varphi(B) > 0$ implies $G(x, B) = \infty$ for all $x \in E$, i.e., when X is weakly recurrent (with respect to φ).

Proof.

See Jain and Jamison (1967, Theorem 2.2). ■

2.8. Existence and Uniqueness of Invariant Measures

In the previous section, we have called a measure m invariant for the Markov chain X (on the state space $(E, \mathcal{B}(E))$) if m satisfies the equation

$$m(B) = \int_E \mu(x, B)\, m(dx)$$

for all $B \in \mathcal{B}(E)$.

The aim of this section is to describe conditions under which such a measure exists and is unique. There is a substantial amount of literature on this subject but, for our purpose, we will quote results taken from Tweedie (1976), Stettner (1988), Orey (1971), Doob (1953), as well as Jain and Jamison (1967). The terms used below are defined in Subsection 2.7.

Consider the following two assumptions:

(A) There exists a finite measure $\varphi \in \mathcal{M}(E)$ such that any invariant set for the process X has positive φ measure.

(B) Any invariant set for the process X is properly essential.

Then the following theorem holds true:

Theorem 1.

Let X be a Markov chain with state space $(E, \mathcal{B}(E))$. Then, under (A) and (B), the following statements are verified:

a) There exists a countable maximal family $\{A_n ; n \in \mathbb{N}\}$ of disjoint φ-minimal invariant subsets of E. (Maximal means that if $\{B_n ; n \in \mathbb{N}\}$ is any other such family with $\{A_n ; n \in \mathbb{N}\} \subset \{B_n ; n \in \mathbb{N}\}$, then $\{A_n ; n \in \mathbb{N}\} = \{B_n ; n \in \mathbb{N}\}$.)

b) Any φ-minimal invariant set $A_{n_0} \in \mathcal{B}(E)$ contains a further set I_{n_0} with $\varphi(A_{n_0}) = \varphi(I_{n_0})$ and which satisfies, for all $D \in \mathcal{B}(E)$, the φ-recurrence property

$$\varphi(D \cap I_{n_0}) > 0 \text{ implies } L(x, D) = Q(x, D) = 1$$

for all $x \in I_{n_0}$.

c) With every set I_n is associated a unique (up to a multiplicative constant) σ-finite invariant measure π_n with supp $(\pi_n) = \overline{I_n}$ and such that $\varphi_{|I_n} \ll \pi_n$, where $\varphi_{|I_n}$ denotes the restriction of the measure φ to the set I_n.

d) Any invariant σ-finite measure m on $(E, \mathcal{B}(E))$ is a linear combination of the π_i's, i.e., there exists a sequence $\{\alpha_n ; n \in \mathbb{N}\}$ such that $m = \Sigma\, \alpha_i\, \pi_i$.

Proof.

See Stettner (1988, Theorems 1.1 and 1.2). ∎

Remark 1.

1) In Assumption (A), instead of a finite measure φ, one can require the existence of a σ-finite measure satisfying the same condition. Indeed, given any σ-finite measure $\overline{\varphi}$, one can construct an equivalent finite measure φ by setting, for $B \in \mathcal{B}(E)$, $\varphi(B) = \int_B f(x)\, \overline{\varphi}(dx)$, where $f(x) \equiv 2^{-n} \left[\overline{\varphi}(E_{n+1} \setminus E_n)\right]^{-1}$ for $x \in E_{n+1} \setminus E_n$ $(E_0 = \phi)$, where the sequence of sets $\{E_n ; n \in \mathbb{N}\}$ satisfies $E_n \in \mathcal{B}(E)$, $\overline{\varphi}(E_n) < \infty$, $\bigcup \{E_n ; n \in \mathbb{N}\} = E$, and $E_n \subset E_{n+1}$.

2) The recurrence property satisfied by any set I_{n_0} is nothing but the usual Harris condition also referred to as strong recurrence by Tweedie (1976).

3) As pointed out by Stettner (1988, Remark 1.2), Condition (B) is used to avoid situations similar to the deterministic motion to the right on $\mathbb{N}$,

$$P(x_{n+1} = x+1 \mid x_n = x) = 1.$$

Indeed, in this case, $\mathbb{N}$ and $E_n = \mathbb{N} \setminus \{0, 1, ..., n\}$ are invariant sets but $\bigcap_{n=0}^{\infty} E_n = \phi$, which prevents the existence of an invariant probability measure (the counting measure is a σ-finite invariant measure for this process).

More specifically, improperly essential sets are the countable union $\{B_n ; n \in \mathbb{N}\}$ of strongly transient sets (see Jain and Jamison (1967, Corollary 4.1)), i.e., $G(x, B_n) < \infty$ for all $x \in E$ and all $n \in \mathbb{N}$. Hence, any σ-finite (sub)invariant measure λ must satisfy $\lambda(E) = \infty$ (see Jain and Jamison (1967, Corollary 4.3)).

Another answer to the aforementioned problem is given in the following result. Again, see Subsection 2.7 for a definition of the terms used.

<u>Theorem 2.</u>

Let μ be the transition kernel of a Markov chain X with state space $(E, \mathcal{B}(E))$. Assume there exists a probability measure φ on $(E, \mathcal{B}(E))$ such that X is φ-irreducible and weakly recurrent, i.e., for all $B \in \mathcal{B}(E)$,

$$\varphi(B) > 0 \text{ implies } \sum_{n=1}^{\infty} \mu(n, x, B) = \infty \quad \text{for all } x \in E.$$

Then

a) There exists a unique nontrivial σ-finite invariant measure λ with $\varphi << \lambda$, and λ is equivalent to the maximal irreducibility measure $\tilde{\varphi}$ for X.

b) The chain X is positive if and only if

$$\int_B R(x, B)\, \lambda(dx) < \infty$$

for one, and then every, $B \in \mathcal{B}(E)$ such that $\varphi(B) > 0$.

c) If X is positive (ergodic), then $\lambda(E) < \infty$ and the unique invariant measure λ (after being normalized to a probability measure) satisfies

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \mu(k, x, B) = \lambda(B) \qquad \tilde{\varphi}\text{-a.e.},$$

for all $B \in \mathcal{B}(E)$.

d) If X is positive and strongly recurrent (i.e., Harris or φ-recurrent), then the $\tilde{\varphi}$-null set over which the limit in (c) above may fail is empty.

Proof.

See Tweedie (1976, Propositions 4.2 and 4.3) as well as, for (a), Stettner (1988, Theorem 2.1). ■

Remark 2.

Proposition 3.6 in Tweedie (1976) states that, under separability, a φ-irreducible and weakly recurrent chain can be restricted to a strongly recurrent chain by removing a $\tilde{\varphi}$-null set. In the same proposition, it is also shown that strong irreducibility and weak recurrence imply strong recurrence, i.e., under ergodicity, strong irreducibility suffices to ensure that the null set in Theorem 2 (c) is in fact an empty set.

Remark 3.

1) If the positivity requirement is omitted in parts (c) and (d) of the above theorem, then we still have a unique invariant measure but it is not necessarily a finite

measure. In fact, one often uses the term ergodic to qualify a chain which possesses a unique finite invariant measure (see, e.g., Tweedie (1975, p. 386)). Note that the existence of such a unique finite invariant measure for a Markov chain X guarantees the positivity of X by Theorem 2 (c) and Tweedie (1976, Proposition 4.2 (i)). Moreover, lack of positivity yields that there exists a sequence of set $B_i \in \mathcal{B}(E)$ such that $B_i \uparrow E$ with each B_i being strongly null, i.e., for each i, $\lim_{n\to\infty} \mu(n, x, B_i) = 0$ for all $x \in E$ (Tweedie (1976, Proposition 4.2.)).

2) Under the conditions of part (d) of Theorem 2 (ergodicity and strong recurrence), we have the stronger (since it is a uniform result over all $B \in \mathcal{B}(E)$) statement that, for any initial distribution ν on $\mathcal{B}(E)$ and with $\|.\|$ denoting the total variation,

$$\lim_{n\to\infty} \left\| \frac{1}{n} \int_E \nu(dy) \sum_{i=1}^{n} \mu(i, y, .) - \lambda(.) \right\| = 0.$$

This result is due to Orey (1971, Proposition 6.1).

Theorem 3.

Suppose the Markov chain X is ergodic and strongly recurrent so that a unique finite invariant measure λ exists. Then, upon normalization of λ to a probability measure, the following statements are verified:

a) If the unique invariant set cannot be decomposed into a cycle (aperiodic case), then the Ceasaro limit $\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mu(k, x, B) = \lambda(B)$ of Theorem 2 (d) is a simple limit. In fact, for all initial probability distribution ν on $\mathcal{B}(E)$, we have

$$\lim_{n\to\infty} \left\| \int_E \mu(n, y, .)\, \nu(dy) - \lambda(.) \right\| = 0,$$

which implies that

$$\lim_{n\to\infty} \mu(n, x, B) = \lambda(B),$$

for all $B \in \mathcal{B}(E)$ and all $x \in E$.

Moreover, we have that, for all $B \in \mathcal{B}^{\infty}(E)$ (the Borel σ-algebra on $E^{\mathbb{N}}$) and all initial distributions ν on $\mathcal{B}(E)$,

$$P_\nu((X_n, X_{n+1}, \ldots) \in B) \longrightarrow P_\lambda((X_1, X_2, \ldots) \in B) \text{ as } n \to \infty,$$

where P_ν and P_λ denote probabilities obtained when $X_0 \sim \nu$ and $X_0 \sim \lambda$ respectively, $\nu, \lambda \in \mathcal{P}(E)$. This implies that the process X with initial distribution ν converges weakly to the unique stationary and ergodic process X with initial distribution λ.

b) If the unique invariant set can be decomposed into a cycle $\{C_n ; 1 \le n \le d\}$, $d > 1$, then there exists a unique collection of probability measures $\{\pi_n ; 1 \le n \le d\}$ with $\pi_i(C_k) = 0$ for $i \neq k$, $\pi_i(C_i) = 1$, and such that, for all $B \in \mathcal{B}(E)$ and all $x \in E$,

$$\lim_{n\to\infty} \mu(nd+m, x, B) = \pi_k(B), \quad x \in C_i \text{ and } k = i + m \pmod d.$$

Moreover, for all $B \in \mathcal{B}(E)$ and all $x \in E$, $\lambda(B)$ above is given by

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mu(k, x, B) = d^{-1} \sum_{n=1}^{d} \pi_n(B),$$

and, for all initial distribution ν on $\mathcal{B}(E)$, we have

$$\lim_{n\to\infty} \left\| \frac{1}{d} \int_E \sum_{k=1}^{d} \mu(n+k, y, .)\, \nu(dy) - \lambda(.) \right\| = 0.$$

Proof.

a) See Orey (1971, Theorem 7.1).

The weak convergence statement comes from a trivial extension of Proposition 7.12 in Breiman (1968).

b) See Orey (1971, Theorem 7.1) as well as Doob (1953, pp. 205–211) and Jain and Jamison (1967, Theorem 2.4). ■

Remark 4.

In the above theorem, it is necessary to require the process to be strongly recurrent instead of simply weakly recurrent in order to obtain the weak convergence result $\lim_{n\to\infty} \mu(n, x, B) = \lambda(B)$ for all $B \in \mathcal{B}(E)$ and all $x \in E$. Failing to do so would give us convergence only if the initial probability measure ν does not put any mass on points x in the $\tilde{\varphi}$-null set over which this limit statement is incorrect and from which the process X can get trapped within sets of $\tilde{\varphi}$ measure zero. By Theorem 2 (d), strong recurrence prevents this from happening.

2.9. Lyapunov Exponents and Oseledec̆'s Multiplicative Ergodic Theorem

In this final introductory section, we will review the basic material related to the theory of Lyapunov exponents and, in particular, state Oseledeč's Multiplicative Ergodic Theorem. This entire discussion is essentially based on Arnold and

Wihstutz's (1986) survey paper. For the main theorem of this section, the reader may also wish to refer to Oseledeč's (1968) original paper.

A review of the theory of Lyapunov exponents can suitably be initiated by some results concerning the stability at the origin of the linear and homogeneous differential equations of the form

$$\dot{x}(t) = A(t)\, x(t)$$

with $x(0) = x_0 \in \mathbb{R}^d$, $t \in \mathbb{R}^+$, and $A : \mathbb{R}^+ \longrightarrow Gl(d, \mathbb{R})$ bounded and continuous.

There are several definitions which relate to the stability of such an equation (see Miller and Michel (1982, Section 5.3)), but the following is the most relevant:

Definition 1.

The equilibrium point $x = 0$ of the above equation is said to be exponentially stable if there exists an $\alpha > 0$ and, for every $\epsilon > 0$, there exists a $\delta_\epsilon > 0$ such that

$$|x(t, t_0, y)| \leq \epsilon \exp[-\alpha (t - t_0)] \quad \text{for all } t \geq t_0,$$

whenever $|y| < \delta_\epsilon$ and $t_0 \geq 0$, and where $x(t, t_0, y)$ represents the solution of the equation which goes through y at time t_0.

It is well known that, when $A(t) = A$ is constant, the stability behavior of this equation is determined by the real part of the eigenvalues of A. Specifically, the differential equation $\dot{x}(t) = A\, x(t)$ is exponentially stable (at the origin and hence everywhere by the nature of the solution of such an equation) if and only if all of the

eigenvalues of A have negative real parts (see, e.g., Miller and Michel (1982, Theorem 5.5.5)).

If the matrix A (t) is periodic, the statements of the previous paragraph remain true if we use the real part of the characteristic exponents of A (t) (see, e.g., Miller and Michel (1982, Theorem 5.5.7)).

Now it is also known that the exponential stability of the above system with A (t) constant or periodic also implies the exponential stability of the associated perturbed nonhomogeneous equation

$$\dot{x}(t) = A(t)\,x(t) + f(t, x(t)),$$

where $f : \mathbb{R}^+ \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$ satisfies $f(t, x(t)) = o(|x|)$ as $|x| \longrightarrow 0$ (see Miller and Michel (1982, Corollary 6.2.5)).

Lyapunov exponents which we now define can be viewed as a generalization of the real parts of the eigenvalues of A (or of the characteristic exponents of A (t)).

Definition 2.

The Lyapunov exponent of a solution $x(t, 0, x_0) \equiv x(t, x_0)$ of the differential equation $\dot{x} = A(t)\,x$ is defined by

$$\lambda(x_0) = \overline{\lim_{t \to \infty}} \frac{1}{t} \log |x(t, x_0)|.$$

Lyapunov proved the following facts about Lyapunov exponents:

1) for all $x_0 \neq 0$, $\lambda(x_0)$ is finite,

2) the set of all numbers which are possible Lyapunov exponents for some initial value x_0 of the differential equation $\dot{x} = A(t)\,x$ is finite, with cardinality p, $1 \leq p \leq d$,

3) for $c \neq 0$, $\lambda(c\,x_0) = \lambda(x_0)$, and

4) for all $c_1, c_2 \in \mathbb{R}$, $\lambda(c_1 x_1 + c_2 x_2) \leq \max\{\lambda(x_1), \lambda(x_2)\}$.

Moreover, if the Lyapunov exponents are ordered by $\lambda_p < \ldots < \lambda_1$, the subspaces $L_i \equiv \{x \in \mathbb{R}^d : \lambda(x) \leq \lambda_i\}$ form a filtration of $\mathbb{R}^d$

$$\phi = L_{p+1} \subset L_p \subset \ldots \subset L_1 = \mathbb{R}^d,$$

with $\dim(L_i) = k_i$ satisfying

$$0 = k_{p+1} < k_p < \ldots < k_1 = d,$$

and

$$\lambda(x) = \lambda_i \text{ if and only if } x \in L_i \setminus L_{i+1},$$

for $1 \leq i \leq p$.

Definition 3.

Let $\{e_1, \ldots, e_d\}$ form a basis of $\mathbb{R}^d$. Then this basis is said to be normal (for the differential equation $\dot{x} = A(t)\,x$) if, for all $c_1, \ldots, c_d \in \mathbb{R}$,

$$\lambda\left[\sum_{i=1}^{d} c_i\, e_i\right] = \max\{\lambda(e_i)\,;\, i \text{ such that } c_i \neq 0\}.$$

Remark 1.

Lyapunov proved that a normal basis always exists.

Let $\{e_1, \ldots, e_d\}$ be a normal basis of $\mathbb{R}^d$ and denote by d_i the multiplicity of λ_i among the numbers $\lambda(e_i)$. Then d_i is the same for any normal basis and

$$d_i = k_i - k_{i+1} \text{ with } \sum_{i=1}^{d} d_i = d.$$

Definition 4.

The Lyapunov exponents λ_i together with their multiplicities d_i are called the Lyapunov spectrum of $\dot{x} = A(t)\,x$.

From the definition of exponential stability and of Lyapunov exponents for $\dot{x} = A(t)\,x$, it should be obvious that this differential equation is exponentially stable if and only if $\lambda_1 < 0$. Unfortunately, it is in general not true that $\lambda_1 < 0$ implies the exponential stability of the associated perturbed nonlinear differential equation $\dot{x} = A(t)\,x + f(t,x)$. For this, one needs that the Lyapunov exponents satisfy a regularity property.

Definition 5.

The differential equation $\dot{x} = A(t)\,x$ (and its associated Lyapunov spectrum) is said to be forward regular if

$$\sum_{i=1}^{p} d_i \lambda_i = \lim_{t\to\infty} \frac{1}{t} \log \det(\phi(t)) = \lim_{t\to\infty} \frac{1}{t} \int_0^t \operatorname{trace}(A(s))\, ds,$$

where $\phi(t)$ is the fundamental matrix of the differential equation with $\phi(0) = \mathrm{Id}$.

For a forward regular system, all the $\overline{\lim}$'s are actually limits and $\lambda_1 < 0$ implies stability of the perturbed system $\dot{x} = A(t)x + f(t,x)$.

To pursue this discussion, we now assume that the differential equation $\dot{x} = A(t)x$ is defined on the whole time axis $\mathbb{R}$ (instead of $\mathbb{R}^+$). Then we can define (backward) Lyapunov exponents by

$$\lambda^-(x_0) = \overline{\lim_{t \to -\infty}} \frac{1}{|t|} \log |x(t, x_0)|$$

and there will be p^-, $1 \leq p^- \leq d$, exponents $\lambda_{p^-}^- < \ldots < \lambda_1^-$ with multiplicities d_i^- for λ_i^-, and a filtration $\phi = L_{p^-+1}^- \subset L_{p^-}^- \subset \ldots \subset L_1^- = \mathbb{R}^d$ with $\dim(L_i^-) = k_i^-$ satisfying $0 = k_{p^-+1}^- < k_{p^-}^- < \ldots < k_1^- = d$, and $\lambda^-(x) = \lambda_i^-$ if and only if $x \in L_i^- \setminus L_{i+1}^-$, for $1 \leq i \leq p^-$.

The differential equation $\dot{x} = A(t)x$, $t \in \mathbb{R}$, (and its associated Lyapunov spectrum) is called backward regular if

$$\sum_{i=1}^{p^-} d_i^- \lambda_i^- = \lim_{t \to -\infty} \frac{1}{|t|} \log \det(\phi(t)) = \lim_{t \to -\infty} \frac{1}{|t|} \int_0^t \operatorname{trace}(A(s))\, ds.$$

In general, the forward and backward spectrum of $\dot{x} = A(t)x$, $t \in \mathbb{R}$, are not related but we have the following definition:

Definition 6.

The differential equation $\dot{x} = A(t)\,x$, $t \in \mathbb{R}$, (and its associated Lyapunov spectrum) is said to be regular if

1) it is forward and backward regular,

2) $p^- = p$ and, for $1 \le i \le p$, $d_i^- = d_{p+1-i}$, $\lambda_i = -\lambda_{p+1-i}$, and

3) for $1 \le i \le p-1$, $L_{i+1} \cap L^-_{p+1-i} = \phi$.

Given a regular system, the subspaces $E_i \equiv L_i \cap L^-_{p+1-i}$, $1 \le i \le p$, have $\dim(E_i) = d_i$ and form a splitting of $\mathbb{R}^d$ according to

$$L_i = \bigoplus_{j=i}^{p} E_j \text{ and } \mathbb{R}^d = \bigoplus_{j=1}^{p} E_j,$$

for $1 \le i \le p$. Moreover,

$$\lim_{t \to \pm\infty} \frac{1}{t} \log |x(t, x_0)| = \lambda_i \text{ if and only if } x_0 \in E_i \setminus \{0\}.$$

Remark 2.

If $A(t)$ is constant or periodic, the system $\dot{x} = A(t)\,x$, $t \in \mathbb{R}$, is regular. In this case, the Lyapunov exponents are simply the real parts of the eigenvalues of A or the real parts of the characteristic exponents of $A(t)$ while the E_i's are the (generalized) eigenspaces.

The fundamental matrix $\phi(t, A)$ associated with the differential equation $\dot{x} = A(t)\,x$ is a $Gl(d, \mathbb{R})$ valued function which, for all s, $t \in \mathbb{R}$, satisfies the relation

$$\phi(t+s, A) = \phi(t, H_s A)\, \phi(s, A).$$

This equation says that ϕ is a multiplicative cocycle (defined below) associated with the dynamical system consisting of the group of shifts $\{H_t ; t \in \mathbb{R}\}$ (in the space of continuous functions $A : \mathbb{R} \longrightarrow gl(d, \mathbb{R})$) defined by $H_t(A(s)) = A(s+t)$.

Since the definition of (forward) Lyapunov exponents could be rewritten as

$$\lambda(A, x_0) = \overline{\lim_{t \to \infty}} \frac{1}{t} \log |\phi(t, A)\, x_0|,$$

(and similarly for backward exponents), Lyapunov exponents are seen to simply describe the asymptotic behavior of the multiplicative cocycle induced by the map A, applied to $x_0 \neq 0$. This is the starting point for the generalization of the above concepts.

Definition 7.

Let $(E, \mathcal{F})$ be a measurable space. We define a measurable flow to be a mapping $H : T \times E \longrightarrow E$, $T = \mathbb{Z}$ or $\mathbb{R}$, such that $\{H_t = H(t, .) ; t \in T\}$ is a collection of bi–measurable bijections from E to E (i.e., H_t and H_t^{-1} are measurable) satisfying

1) for all measurable maps $f : E \longrightarrow \mathbb{R}$, $f(H_t x)$ is measurable in (t, x) and

2) $\{H_t\}$ satisfies the group or flow property, i.e., for all $s, t \in T$,

$$H_{t+s} = H_t \circ H_s.$$

Remark 3.

H_t is typically a shift on the space of functions determining the dynamics of the system. In the stochastic case, E is the underlying probability space, i.e., in our

case, the path space of the stochastic process and, for control systems, E is the space of admissible control functions $\mathcal{U}$.

The solution of differential or difference equations describing the systems dynamics induces a "flow" on the state space M (a C^∞ manifold), which is intertwined with H in the form of a skew product flow:

Definition 8.

A skew product flow on $E \times M$ is a map

$$\varphi : T \times (E \times M) \longrightarrow (E \times M)$$

defined by

$$\varphi(t, x, m) = (H_t\, x, F(t, x, m)),$$

where $H : T \times E \longrightarrow E$ is a measurable flow and $F : T \times E \times M \longrightarrow M$ satisfies the skew product property

$$F(t+s, x) = F(t, H_s\, x) \circ F(s, x).$$

Here we have used the notation $F(t, x) : M \longrightarrow M$ for the component map $F(t, x, .)$ for all $(t, x) \in T \times E$.

Note that φ as defined above is a flow in the sense of Definition 7 since

$$\begin{aligned}\varphi(t+s, x, m) &= (H_{t+s}\, x, F(t+s, m, x) \\ &= (H_t\,(H_s\, x), F(t, H_s\, x, F(s, x, m))) \\ &= \varphi(t, H_s\, x, F(s, x, m))\end{aligned}$$

$$= \varphi(t, \varphi(s, x, m)),$$

which, using as above $\varphi_t : E \times M \longrightarrow M$ for the component map $\varphi(t, ., .)$ for all $t \in T$, can be written $\varphi_t \circ \varphi_s(x, m)$. But, for fixed $x \in E$, $F_x : T \times M \longrightarrow M$ is not a flow.

Remark 4.

For deterministic dynamical systems, whose dynamics do not explicitly depend on $t \in T$, one does not need the shift space (E, H). Consider, e.g., the differential equation $\dot{x} = f(x)$ on M. Here the flow φ takes the form $\varphi : T \times M \longrightarrow M$ and $\varphi(t, m)$ denotes the solution at time $t \in T$ with initial value $\varphi(0, m) = m$.

For time dependent differential equations, the situation is different. Consider, for example, $\dot{x} = A(t)\, x$ in $\mathbb{R}^d$, with $A : \mathbb{R} \longrightarrow gl(d, \mathbb{R})$ bounded continuous. Here we have $E = C(\mathbb{R}, gl(d, \mathbb{R}))$, the continuous functions from $\mathbb{R}$ into $gl(d, \mathbb{R})$, $H : T \times E \longrightarrow E$ is defined by $H(t, A(.)) = A(t+.)$, the usual shift, and $\varphi : T \times E \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$ is given by $\varphi(t, A(.), m) = (H_t\, A(.), F(t, A(.), m))$ where $F(t, A(.), m) = \phi(t, A(.))\, m$, with $\phi(t, A(.))$ representing the fundamental matrix of $\dot{x} = A(t)\, x$ (using $\phi(0, A(.)) = Id$, the $d \times d$ identity matrix).

The situation for control systems is similar except that $E = C(\mathbb{R}, gl(d, \mathbb{R}))$ is replaced by the space $\mathcal{U}$ of admissible control functions.

Finally, for stochastic systems, we consider the case of products of random matrices (this case arising from linear stochastic difference equations on $\mathbb{R}^d_0$ of the form $x_{n+1} = A_n\, x_n$ with A_n being a $Gl(d, \mathbb{R})$–valued random variable). Take $T = \mathbb{Z}$ and $E = Gl(d, \mathbb{R})^{\mathbb{Z}}$, the path space of the stochastic process, with its canonical σ-algebra $\mathcal{F}$, and the usual shift H. Let $A : E \longrightarrow Gl(d, \mathbb{R})$ be measurable and define the

products

$$C(n,x) = \begin{cases} A(H_{n-1}\,x) \ldots A(x) & \text{for } n \geq 1, \\ Id & \text{for } n = 0, \\ A^{-1}(H_n\,x) \ldots A^{-1}(H_{-1}\,x) & \text{for } n \leq -1. \end{cases}$$

Now the sequences $\{A_n \equiv A(H_n\,x)\,;\, n \geq 0\}$ and $\{A_n^{-1} \equiv A^{-1}(H_n\,x)\,;\, n \leq 0\}$ are sequences of $Gl(d,\mathbb{R})$–valued random variables. They induce a skew product flow on $E \times \mathbb{R}_0^d$ through $\varphi : T \times E \times \mathbb{R}_0^d \longrightarrow E \times \mathbb{R}_0^d$ defined by $\varphi(n,x,m) = (H_n\,x, A_n \ldots A_1\,m)$. Note that, if the sequence $\{A_n\,;\, n \in \mathbb{Z}\}$ is stationary, then there exists a probability measure μ on $(E, \mathcal{F})$ which is invariant under the flow H, i.e., $\mu H_n = \mu$ for all $n \in \mathbb{Z}$. This is actually the typical situation in Oseledeč's Theorem (see Theorem 1 later in this section and the discussion thereafter). In particular, the setup described here was studied in detail by Furstenberg and Kesten (1960).

The (exponential) growth behavior of a flow is measured through its Lyapunov exponent, i.e., the growth rate of its linearization. The purpose of such a linearization is to allow the use of the norm on $T_m(M)$ arising from C^∞ Riemannian metric associated with M (see Definition 2.3.9 and Remark 2.3.4). Linearizing a flow means associating a multiplicative cocycle with it. This justifies the next definition:

Definition 9.

Given a measurable flow $\varphi : T \times E \times M \longrightarrow E \times M$, a (multiplicative) cocycle associated with φ is a measurable map

$$C : T \times E \times M \longrightarrow L(T_m(M), T_{F(t,x,m)}(M)) \cong Gl(d,\mathbb{R}),$$

where $L(T_m(M), T_{F(t,x,m)}(M))$ denotes the bijective linear maps from $T_m(M)$ to $T_{F(t,x,m)}(M)$ (identified with the nonsingular $d \times d$ matrices), which, for all s, $t \in T$ and all $m \in M$, satisfies

$$C(t+s, x, m) = C(t, H_s x, F(s, x, m)) \circ C(s, x, m).$$

In particular, $C(t, x, m) : T_m(M) \longrightarrow T_{F(t,x,m)}(M)$ is a vector space isomorphism.

Recall that $T(M)$ was defined to be the tangent bundle of the manifold M (see Subsection 2.3). The elements in $T(M)$ are represented by the pairs (m, v) where $m \in M$ and v is a tangent vector in $T_m(M)$. Combining the flow φ and its associated cocycle C, one obtains a (linearized) skew product flow $\tilde{\varphi} : T \times E \times T(M) \longrightarrow E \times T(M)$ defined by

$$\tilde{\varphi}(t, x, (m, v)) = (H_t x, F(t, x, m), C(t, x, m)\, v),$$

where $v \in T_m(M)$.

The question now is about how one does assign a cocycle to a given flow, i.e., about how one does linearize a given flow φ. For a linear system in $\mathbb{R}^d$, we can, for all $m \in \mathbb{R}^d$, identify $T_m(\mathbb{R}^d)$ with $\mathbb{R}^d$ itself and the linearization of a linear system yields the same flow. This applies to the last two examples in Remark 4.

To be more precise, consider again the differential equation $\dot{x} = A(t)\, x$, $E = C(\mathbb{R}, gl(d, \mathbb{R}))$. One can see that, in the definition of the skew product flow φ, the component $F(t, A(.), m)$ describes the effect at time t of the flow φ on the point

m while $C(t, A(.), m)\, v$ describes the corresponding effect at time t of the linearized flow on the tangent vector $v \in T_m(M)$. Hence $C(t, A(.), m)\, v$ solves the equation $x\left[\frac{d}{dt}\right] = F_*(t, A(.), .)$ at $v \in T_m(\mathbb{R}^d)$. In the above linear situation and upon identification of $T_m(\mathbb{R}^d)$ with $\mathbb{R}^d$, $F_*(t, A(.), m)$ turns out to be simply the Jacobian of $A(t)\, x$ at $m \in \mathbb{R}^d$. Since the fundamental matrix $\phi(t, A(.))$ associated with this differential equation will satisfy (with $v \in \mathbb{R}^d \cong T_m(\mathbb{R}^d)$ for all m)

$$\phi(t, A(.))\, v = F(t, A(.), m)\, v$$

and since the equation $x\left[\frac{d}{dt}\right] = F_*(t, A(.), .)$ describes the same dynamics as the original equation $\dot{x} = A(t)\, x$, we see that, for this (and, in general, for all) linear system, $C(t, A(.), m) = F(t, A(.), m)$.

Similarly, for the sequence of random variables $\{A_n \equiv A(H_n\, x)\, ;\, n \in \mathbb{Z}\}$, the cocycle is the product of these matrices $C(t, x)$, as defined in Remark 4. Note that, again because of the identification of $T_m(\mathbb{R}^d)$ with $\mathbb{R}^d$, the argument m in this cocycle notation is unnecessary and was omitted.

For nonlinear systems, we have to distinguish between the continuous time ($T = \mathbb{R}$) and the discrete time ($T = \mathbb{Z}$) cases. In the continuous time case, systems are given through vector fields on M and the linearized system is described locally by the Jacobian matrix of the vector field. Similarly to the linear case discussed above, the cocycle $C(t, x, m)$ is then (locally) the fundamental matrix of the linearized system (see, for example, Kliemann (1988, Section 3.1), for the precise description in different situations).

In the discrete time case, systems are given via difference equations $y_{n+1} = f(y_n, H_n x)$, $y \in M$, $x \in E$, with $f \in \mathrm{Diff}(M)$ and $H_n x = x_n$. Then the linearization, at the n^{th} step, is the map

$$f_*^n \equiv (f(y_n, H_n x))_* : T_{y_n}(M) \longrightarrow T_{f(y_n, H_n x)}(M),$$

as defined in Subsection 2.3. (Remember that, if $f \in \mathrm{Diff}(M)$, f_*^n is a vector space isomorphism.) The cocycle is the product of these maps, i.e.,

$$C(n, x, y) = (f_{H_{n-1} x \ldots H_0 x}(y))_* : T_y(M) \longrightarrow T_{f_{H_{n-1} x \ldots H_0 x}(y)}(M),$$

with $(f_{H_{n-1} x \ldots H_0 x}(y))_* = f_*^{n-1} \ldots f_*^0$.

If $y_{n+1} = f(y_n)$, $C(n, x, y) = C(n, y) = (f)_*^n : T_y(M) \longrightarrow T_{(f)^n(y)}(M)$, with $(f)^n$ denoting the n–fold composition of the map f. Mañé (1987) writes $D_y f^n$ for the linearization $(f)_*^n$ of the iterated map $(f)^n$. A detailed account of the ergodic and entropy theory for such systems (and much more) can be found in his book (see, in particular, Sections 4.10 and 4.11).

Definition 10.

Given the linearized flow $\tilde{\varphi} : T \times E \times T(M) \longrightarrow E \times T(M)$, where M is a C^∞ Riemannian manifold, the (forward) Lyapunov exponent of $\tilde{\varphi}$ is defined by

$$\lambda(x, m, v) = \overline{\lim_{t \to \infty}} \frac{1}{t} \log \| C(t, x, m)\, v \|,$$

where $\|.\|$ is the norm associated with the Riemannian metric on $T(M)$. In other words, $\lambda(x, m, v)$ is the exponential growth of the cocycle at (x, m) in the direction of $v \in T_m(M)$.

Under the condition $\|C(t,x,m)\| \leq \exp(|t|\, c(x,m))$ as $|t| \longrightarrow \infty$ for some measurable $c : E \times M \longrightarrow \mathbb{R}$ ($\|.\|$ is the operator norm associated with the Riemannian metric on $T(M)$), $\lambda(x,m,v)$ is finite for all (x,m,v) provided $v \neq 0$. Moreover, all the statements made for the forward exponents of the system $\dot{x} = A(t)\,x$, $t \in \mathbb{R}$, remain true for $\lambda(x,m,v)$.

Since $\lambda(x,m,v)$ is measurable in all variables, the functions $p(x,m)$, $\lambda_i(x,m)$, $L_i(x,m)$, $k_i(x,m)$, and $d_i(x,m)$ are also measurable. Moreover the equation

$$\tilde{\varphi}_t\, \lambda(x,m,v) \equiv \lambda(H_t\, x, F(t,x,m), C(t,x,m)\, v) = \lambda(x,m,v)$$

shows that $\lambda(x,m,v)$ is $\{\tilde{\varphi}_t\}$ – invariant and therefore,

$$H_t\, p(x,m) = p(x,m),$$

$$H_t\, \lambda_i(x,m) = \lambda_i(x,m),$$

$$H_t\, k_i(x,m) = k_i(x,m),$$

$$H_t\, d_i(x,m) = d_i(x,m), \text{ and}$$

$$H_t\, L_i(x,m) \equiv L_i(H_t\, x, F(t,x,m)) = C(t,x,m)\, L_i(x,m).$$

The definition of backward exponents and regular cocycle C (or regular skew product flow) carries over from the definitions at the beginning of this section, simply replacing the specific cocycle $\phi(t,A)$ there by $C(t,x,m)$. The regularity of such a cocycle will depend on $(x,m) \in E \times M$. For every regular point (x,m), the limit superior is in fact a limit, the subspaces $E_i(x,m)$ resulting from the splitting of $T(M)$ described before will satisfy

$$E_i\,(H_t\,x, F\,(t,x,m)) = C\,(t,x,m)\,E_i\,(x,m),$$

and again

$$\lim_{t\to\pm\infty} \frac{1}{t}\log\,|C\,(t,x,m)\,v| = \lambda_i\,(x,m) \text{ if and only if } v \in E_i\,(x,m) \setminus \{0\}.$$

Now the question is to find conditions which will ensure regularity of a point $(x,m) \in E \times M$. The crucial assumption under which this problem can be dealt with is the existence of a probability measure μ on $E \times M$ which is invariant with respect to the flow $\{\varphi_t\,;\, t \in T\}$, i.e., $\varphi_t\,\mu = \mu$ for all $t \in T$. Under this hypothesis, we have the following key theorem:

Theorem 1. (Oseledeč's Multiplicative Ergodic Theorem.)
Let $\{\tilde{\varphi}_t\}$ be a skew product flow on $E \times T\,(M)$ and assume that there is a measure μ on $E \times M$ which is left invariant by the skew product flow $\{\varphi_t\}$. Also assume that

$$\int_{E\times M} \sup_{-1\le t\le 1} \log^+ \|C^{\pm 1}\,(t,x,m)\|\,d\,\mu\,(x,m) < \infty,$$

where $\log^+(a) = \max\,(0, \log(a))$.

Then there is a measurable set $\Gamma \subset E \times M$ with $\varphi_t\,\Gamma = \Gamma$ and $\mu\,(\Gamma) = 1$ such that:

1) every $(x,m) \in \Gamma$ is regular (so that all results for regular points apply μ-a.s.) and

2) if μ is ergodic for φ_t, i.e., if every φ_t invariant set has μ measure 0 or 1, then p, λ_i, k_i, and d_i are constants (but $L_i\,(x,m)$ and $E_i\,(x,m)$ still depend on (x,m)).

Proof.

This result is an abridged version of the Theorem found p. 8 in Arnold and

Wihstutz (1986). Also see Kliemann (1988, Theorem 3.2.4) and Oseledeč's (1968) original paper. ■

Recall that, for already linear time dependent situations on $\mathbb{R}^d$, there is no need to linearize the skew product flow φ and this skew product flow can be directly used in Oseledeč's Theorem (i.e., the skew product flow $\tilde{\varphi}$ above reduces to φ while the flow leaving the measure μ left invariant is simply H (on E)). In this case, $E \times T(M)$ reduces to $E \times \mathbb{R}^d$ (using $T_m(\mathbb{R}^d) \cong \mathbb{R}^d$ for all $m \in \mathbb{R}^d$) and, as explained before, the cocycle $C(t, x, m)$ is then simply $F(t, x, m)$. This setup is basically the formulation of Oseledeč's Theorem found in Arnold and Wihstutz (1986).

An example of this type, to which Oseledeč's Theorem can be applied, was presented in Remark 4. If the sequence of random matrices $\{A_n ; n \in \mathbb{Z}\}$ is stationary, then there is a H_n invariant probability measure μ. The integrability condition in Theorem 1 reads for this case:

$$\log^+ \|A(.)\| \text{ and } \log^+ \|A^{-1}(.)\| \in L^1(E, \mathcal{F}, \mu).$$

Under this condition, the linear flow φ on $E \times \mathbb{R}^d$ has regular elements in E with μ probability one. A variety of other examples are discussed in Kliemann (1988) for the continuous time case. Further research on the linear, discrete time case is the topic of this thesis.

3. STUDY OF CONTROL SETS FOR DETERMINISTIC SYSTEMS

3.1. Control Sets of Semigroups Associated with Discrete Dynamical Systems

In subsequent sections, we aim to investigate the behavior of stochastic discrete time dynamical systems. This will be done using the tools of control theory, especially via the notion of control sets. We therefore need to discuss the basic features of such control sets.

In the following, M^d will denote a C^∞ connected Riemaniann manifold with metric ϱ as discussed in Subsection 2.3. We will write Diff (M) for the collection of all diffeomorphisms from M to M. Under the composition law, Diff (M) is clearly a group and, using the open–compact topology, it is a topological group (refer to Subsections 2.1 and 2.2). $\mathcal{S}$ and $\mathcal{G}$ will be used to denote, respectively, an arbitrary semigroup and an arbitrary group contained in Diff (M). If $\mathcal{X}$ denotes an arbitrary subset of Diff (M), we will write $\mathcal{S}(\mathcal{X})$ and $\mathcal{G}(\mathcal{X})$ to represent the semigroup and the group generated by $\mathcal{X}$, respectively. In particular, the group generated by the semigroup $\mathcal{S}$ will be denoted by $\mathcal{G}(\mathcal{S})$.

Semigroups (and, whenever defined, groups) are useful in the study of any discrete dynamical control system Σ consisting of:

1) a state space M^d,

2) a set of control values $U \subset \mathbb{R}^k$ with a set of admissible control functions $\mathcal{U} = \{u : \mathbb{Z} \to U\}$, and

3) the dynamics $x_{n+1} = h(x_n, u_n)$ where $\mathcal{H} = \{h(.,u) ; u \in U\}$ is a collection of continuous maps from M to M, indexed by $u \in U$.

4) In many cases, authors use additional requirements.
Meyn and Caines (1988) require that the map $h : M \times U \longrightarrow M$ be C^1 with U being an open subset of $\mathbb{R}^k$ and Meyn (1989) requires h to be C^∞.
Jacubczyk and Sontag (1988) require $\mathcal{H} = \{h(.,u) ; u \in U\} \subset \text{Diff}(M)$ and h to be C^∞ with $U \subset \text{int } \overline{U}$, $0 \in U$.

Indeed, the solution to such a system simply consists of the trajectories or paths starting at the initial point $x_0 \in M$ and associated with $\mathcal{S}(\mathcal{H})$ (or $\mathcal{G}(\mathcal{H})$ if the dynamical system can be run backward in time (if $\mathcal{G}(\mathcal{H})$ is defined)), i.e., of the sequences $h_n \circ \ldots \circ h_1(x_0)$, $h_i \in \mathcal{H}$, $i \in \{1, \ldots, n\}$, $n \in \mathbb{N}$. The orbits associated with $\mathcal{S}(\mathcal{H})$ (see Subsection 2.4) are simply subsets of M and hence contain much less information than the paths or trajectories arising from the dynamics of a system (loosely speaking, the orbit of a point is concerned about where you can go from that point but not about how you can get there). In many instances, results can be proved using only the notion of orbit. This is the case for most results in this section (unless explicitely stated otherwise) but, in Subsection 3.3, we will discuss results which crucially depend on the paths associated with $\mathcal{S}(\mathcal{H})$ (or $\mathcal{G}(\mathcal{H})$) and not only on the orbits.

The additional requirements on the maps $h(.,u)$, $u \in U$, and h, which were described in part (4) of Definition 1, are typically not needed in this section. The requirement that the maps $h(.,u) : M \longrightarrow M$ describing the dynamics of the system be diffeomorphisms is useful in allowing the use of Lie group theory and will be used

later on. Moreover, the use of diffeomorphisms ensures that the group $\mathcal{G}(\mathcal{X})$ is defined. (This is why diffeomorphisms are used in Definition 2.4.4.) In fact, in Subsection 3.1, only two results require more than continuity of the maps h (., u) (Proposition 3.1.6 requires differentiability and Proposition 3.1.10 needs some additional assumptions). Hence, in Subsection 3.1 and except when otherwise specified, we will only assume that the maps h (., u) : M $\longrightarrow$ M are continuous (and, accordingly, that $\mathcal{S}$ denotes a semigroup of continuous functions). Nevertheless, all but one of our (counter)examples will be built around diffeomorphisms to ensure their validity in even quite restrictive setups (Example 2 (a) does not use diffeomorphisms but was nevertheless given for future reference and because of its simplicity).

In order to conform with the works of other authors, we will also always construct our (counter)examples with the set of control values U as an open and connected subset of $\mathbb{R}^k$. But this is not a necessary assumption for any of the proofs in this section (it is even irrelevant in Proposition 2).

Finally, all (counter)examples given, even when solely described via their semigroups, will always arise from some dynamical system, i.e., h : M $\times$ U $\longrightarrow$ M will be continuous. Moreover, whenever the construction of a specific dynamical system generating the semigroup under discussion is not obvious, an example of the former will be given.

Associated with the orbits generated by a system's group or semigroup is the notion of control set which we define following Kliemann (1979), Arnold and Kliemann (1983), and Arnold et al. (1986a).

Definition 1.

Let $\mathcal{S}$ be a semigroup of continuous maps acting on a manifold M. The positive orbit generated by $\mathcal{S}$ starting at $x \in M$ is defined by $\{y \in M : y = g\,x \text{ for some } g \in \mathcal{S}\}$ and denoted by Sx.

A subset C of M is said to be an invariant control set (for $\mathcal{S}$) if $\overline{Sx} = \overline{C}$ for all $x \in C$.

C is a maximal invariant control set (for $\mathcal{S}$) if C is an invariant control set and, whenever $D \subset M$ is such that $C \subset D$ and $\overline{Sx} = \overline{D}$ for all $x \in D$, we have $C = D$.

A subset D of M is control invariant (in short C–invariant) if $Sx \subset D$ for all $x \in D$.

Remark 1.

From the above definition, it is clear that, if C is an invariant control set, any proper subset of C, which is dense in C, will also be an invariant control set but will not be maximal.

Proposition 1.

Let C be an invariant control set. Then

a) C closed implies C maximal,

b) $x \in \overline{C}$ implies $\overline{Sx} \subset \overline{C}$, i.e., $\overline{C}$ is C–invariant,

c) $x \in \overline{C}$ and $\operatorname{int} \overline{Sy} \neq \phi$ for all $y \in \partial C$ imply $\overline{Sx} = \overline{C}$,

d) C maximal and $x \in \partial C$ with $\operatorname{int} \overline{Sx} \neq \phi$ imply $x \in C$.
In particular, if C is maximal and $\operatorname{int} \overline{Sx} \neq \phi$ for all $x \in \partial C$, then C is closed, and

e) to C corresponds exactly one maximal invariant control set C_0 with $C \subset C_0$.
Moreover, $C_0 \subset \overline{C}$ and $C_0 \neq \overline{C}$ implies $\operatorname{int} \overline{Sx} = \phi$ for some $x \in \partial C$.

Proof.

a) Suppose there is an invariant control set D such that $C \subset D$. Then, for all $x \in D$, $\overline{Sx} = \overline{D}$. In particular, for all $x \in C$, $C = \overline{C} = \overline{Sx} = \overline{D}$ and this obviously implies $C = D$.

b) Since the elements of $\mathcal{S}$ are continuous, if there exists $z \in Sx \cap (\overline{C})^c$, then there exists a whole neighborhood of x mapped into the open set $(\overline{C})^c$. Since any neighborhood of x contains points in C, this contradicts the invariance of C. So $Sx \subset \overline{C}$ and hence $\overline{Sx} \subset \overline{C}$.

c) If $x \in C$, $\overline{Sx} = \overline{C}$ by definition. If $x \in \partial C$, int $\overline{Sx} \neq \phi$ and, by (b), $\overline{Sx} \subset \overline{C}$, i.e., $\overline{Sx} \cap \overline{C} \neq \phi$. Since int $\overline{Sx} \subset \overline{C}$, we can pick $y \in \overline{Sx} \cap C$ and $\overline{Sy} = \overline{C}$. Now, for each $g \in \mathcal{S}$ and every ϵ-neighborhood V^{ϵ}_{gy} of g y, $y \in \overline{Sx}$ implies that $g^{-1}(V^{\epsilon}_{gy})$ contains a point $x^{\epsilon}_{g} \in Sx$. Hence, for all $g \in \mathcal{S}$, every ϵ-neighborhood of g y contains a point in Sx, namely $g\, x^{\epsilon}_{g}$. This means that $Sy \subset \overline{Sx}$. Therefore, $\overline{C} = \overline{Sy} \subset \overline{Sx}$ and, consequently, $\overline{Sx} = \overline{C}$.

d) The proof of (c) shows that int $\overline{Sx} \neq \phi$ implies $\overline{Sx} = \overline{C}$. But, by maximality, this implies $x \in C$.

e) That C_0 is unique is trivial by the definition of maximality.

To see that $C_0 \subset \overline{C}$, simply note that if there exists $x \in C_0 \backslash \overline{C}$, then, for all $y \in C \cap C_0 \neq \phi$, $y \in C_0$ implies $x \in \overline{C_0} = \overline{Sy}$ while $y \in C$ implies $\overline{Sy} = \overline{C}$ and $x \notin \overline{Sy}$, a contradiction.

The last claim follows immediately from part (d).

■

Example 1.

This example shows that a maximal invariant control set need not be closed (nor

open) if int $\overline{Sy} = \phi$ for some (or all) $y \in \partial C$ (also see Example 2 (b)).

Let $M = \mathbb{R}_0^2$ and take the semigroup $\mathcal{S} = \{A_{\alpha\beta}; \alpha > 0 \text{ and } \beta > 0\}$ where

$$A_{\alpha\beta} = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} \in G\ell(2, \mathbb{R}).$$

It is easy to observe that, for instance, any point $z = \begin{bmatrix} x \\ y \end{bmatrix}$, $x > 0$, $y > 0$, will have the open upper right quadrant $\{(x, y) \in \mathbb{R}_0^2 : x > 0,\ y > 0\}$ as its orbit while any point $z = \begin{bmatrix} x \\ 0 \end{bmatrix}$, $x > 0$, will have the positive x–axis $\{(x, y) \in \mathbb{R}_0^2 : x > 0,\ y = 0\}$ as its orbit. Using different initial points, it is then clear that $\mathcal{S}$ generates eight maximal invariant control sets: the four open quadrants and the four half lines formed by the two coordinate axes without the origin.

For future reference, here are interesting variations of the above setup:

a) Take $\mathcal{S}$ $(\{A_{\alpha\beta}; \alpha > 0 \text{ and } \beta > 0\})$, where

$$A_{\alpha\beta} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} \in G\ell(2, \mathbb{R}).$$

Then the open set $\{(x, y) \in \mathbb{R}_0^2 : x \cdot y \neq 0\}$ constitutes an open and disconnected maximal invariant control set (the union of the four open quadrants) while the set $\{(x, y) \in \mathbb{R}_0^2 : x \cdot y = 0\}$ is a closed (in $\mathbb{R}_0^2$) and disconnected maximal invariant control set (the union of the two coordinate axes without the origin).

b) Take $\mathcal{S}$ $(\{A_{\alpha\beta}; \alpha > 0 \text{ and } \beta > 0\} \cup \{B\})$, where $A_{\alpha\beta}$ is as above and $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $B \in Gl(2, \mathbb{R})$. Then $\mathcal{S}$ corresponds to all matrices of the form $\begin{bmatrix} \alpha & \gamma \\ 0 & \beta \end{bmatrix}$, where $\alpha > 0$, $\beta > 0$, and $\gamma > 0$.

As above, the points in the upper right open quadrant have this same quadrant as

their orbit and similarly for the points in the lower left open quadrant. Moreover, if $z \in \{(x,y) \in \mathbb{R}_0^2 : x = 0,\ y > 0\}$, then $Sz = \{(x,y) \in \mathbb{R}_0^2 : x > 0,\ y > 0\}$ while $z \in \{(x,y) \in \mathbb{R}_0^2 : x > 0,\ y = 0\}$ implies that $Sz = \{(x,y) \in \mathbb{R}_0^2 : x > 0,\ y = 0\}$, i.e., under $\mathcal{S}$, the positive y-axis has the right open quadrant as orbit while the the positive x-axis is mapped onto itself. Similarly, under $\mathcal{S}$, the points on the negative y-axis have the lower left open quadrant as orbit while the negative x-axis is mapped onto itself. Finally, if $z \in \{(x,y) \in \mathbb{R}_0^2 : x > 0,\ y < 0\}$ then $Sz \cap \{(x,y) \in \mathbb{R}_0^2 : x < 0,\ y < 0\} \neq \phi$ and $z \in \{(x,y) \in \mathbb{R}_0^2 : x < 0,\ y > 0\}$ implies $Sz \cap \{(x,y) \in \mathbb{R}_0^2 : x > 0,\ y > 0\} \neq \phi$.

The situation is depicted below and exhibits four maximal invariant control sets: two <u>neither open nor closed</u> sets

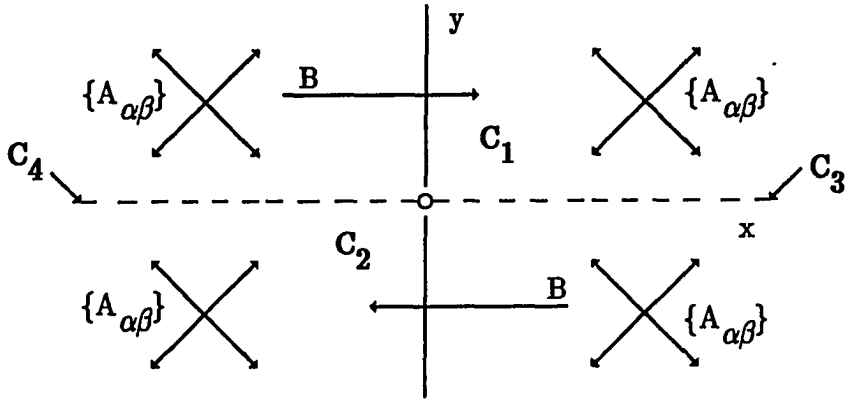
$$C_1 = \{(x,y) \in \mathbb{R}_0^2 : x \geq 0,\ y > 0\} \text{ and}$$

$$C_2 = \{(x,y) \in \mathbb{R}_0^2 : x \leq 0,\ y < 0\},$$

as well as two other sets on the x-axis,

$$C_3 = \{(x,y) \in \mathbb{R}_0^2 : x > 0,\ y = 0\} \text{ and}$$

$$C_4 = \{(x,y) \in \mathbb{R}_0^2 : x < 0,\ y = 0\}.$$



The crossing arrows associated with $\{A_{\alpha\beta}\}$ indicate that the action of the $A_{\alpha\beta}$ matrices will move any point in the indicated quadrant to any other point in the same quadrant. The arrows indexed by B indicate the action of the matrix B allowing points to cross from one quadrant to another.

c) Take $\mathcal{S}(\{A_{\alpha\beta}; \alpha < 0 \text{ and } \beta < 0\})$, where $A_{\alpha\beta}$ is as above. Then $\mathcal{S}$ corresponds to all matrices of the form $\begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix}$, where α and β have the same sign.

After the previous discussions, it is clear that this semigroup generates four disconnected maximal invariant control sets:

$$C_1 = \{(x,y) \in \mathbb{R}_0^2 : x > 0,\ y > 0\} \cup \{(x,y) \in \mathbb{R}_0^2 : x < 0,\ y < 0\},$$

$$C_2 = \{(x,y) \in \mathbb{R}_0^2 : x < 0,\ y > 0\} \cup \{(x,y) \in \mathbb{R}_0^2 : x > 0,\ y < 0\},$$

$$C_3 = \{(x,y) \in \mathbb{R}_0^2 : x \neq 0,\ y = 0\}, \text{ and}$$

$$C_4 = \{(x,y) \in \mathbb{R}_0^2 : x = 0,\ y \neq 0\}.$$

d) Take $\mathcal{S}(\{A_{\alpha\beta}; \alpha > 0 \text{ and } \beta > 0\})$, where $A_{\alpha\beta}$ is as above, and its canonical action on the circle $\mathbb{S}^1$ (i.e., the projection on $\mathbb{S}^1$ of $g\,s$, $g \in \mathcal{S}$, $s \in \mathbb{S}^1$). Identifying the elements of $\mathbb{S}^1$ with their angle in radians, we have again eight maximal invariant control sets: $\{0\}$, $\{\pi/2\}$, $\{\pi\}$, $\{3\pi/2\}$ and the four open sets lying between these points. This provides an example of open maximal invariant control sets on $\mathbb{S}^1$.

Examples on $\mathbb{S}^1$ for neither open nor closed and/or disconnected maximal invariant control sets can easily be constructed from their counterparts in $\mathbb{R}_0^2$.

Also note that all the above semigroups may naturally arise from controlled difference equations of the form $x_{n+1} = h(x_n, u_n)$ as described at the beginning

of this section. Such dynamical systems are quite easy to construct and are not explicitly given here.

Example 2.

The examples below show that, when C is an invariant control set (and hence $\overline{C}$ is C–invariant) or even a maximal invariant control set, it is not necessarily true that either C or int C are C–invariant.

a) Consider the controlled difference equation on $\mathbb{R}$

$$x_{n+1} = a\, x_n,$$

with $a \in \mathbb{R}$.

Then $C = \mathbb{R}_0 = \text{int } C$ is easily seen to be a maximal invariant control set for this system. But $C = \text{int } C$ is not C–invariant since, using the control $a = 0$, one can reach the set $\{0\} = C^c$ (which is itself another maximal invariant control set). Note that in this case the maps $h(x, a) \equiv a\, x$, $a \in \mathbb{R}$, are not diffeomorphisms.

b) Consider the discrete time control system $z_{n+1} = h(z_n, u_n)$ on $\mathbb{R}^2$, where

$$h(z, u) \equiv h((x, y), (\alpha, \beta, \gamma)) = (\beta x + \alpha f(y), \beta x + \gamma y),$$

$(\alpha, \beta, \gamma) \in (-1/4, 1/4) \times \overline{\mathbb{R}_0^-} \times (1, \infty)$, with the map $f : \mathbb{R} \longrightarrow \mathbb{R}$ defined by

$$f(y) = \begin{cases} 0 & \text{for } y \geq 1 \\ \exp\left[-[y-1]^{-2}\right] & \text{for } y < 1. \end{cases}$$

For any $u = (\alpha, \beta, \gamma)$, the map $h(., u)$ can be checked to be a diffeomorphism from $\mathbb{R}^2$ to $\mathbb{R}^2$ (it is one-to-one, onto, with nonzero Jacobian (using $|\alpha| < 1/4$)) and the maximal invariant invariant control set associated with this system can be verified to be

$$C = \mathbb{R}^2 \setminus \{(x,y) \in \mathbb{R}^2 : x = 0, y \geq 1\}.$$

Using $\gamma = 3$ and $\alpha = 0$, one can then see that the point $(0, 1/2) \in \text{int } C$ is mapped to $(0, 3/2) \notin C$.

A related example for the continuous case can be found in Arnold and Kliemann (1987, Remark 3.3.)

Proposition 2.

Let C be an invariant control set. Then either all points in C are isolated (with respect to C) and C is maximal, or no point in C is isolated.

Proof.

Let $x \in C$ be isolated and assume there exists a point $y \in C$, y not isolated, as well as a collection of continuous maps $\{g_n ; n \geq 1\} \subset \mathcal{S}$ such that $g_n\, y \longrightarrow x$.

If $g_{n_0}\, y = x$ for some n_0, pick a neighborhood V_x of x with $V_x \cap C = \{x\}$. Then there is a neighborhood V_y of y such that $g_{n_0}(V_y) \subset V_x$. Since y is not isolated in C, there exists a sequence $\{y_k ; k \geq 1\} \subset V_y \cap C$ with $y_k \neq y$ and $y_k \longrightarrow y$. But then, by the invariance of C, $g_{n_0}\, y_k \subset \overline{C}$ for all k, $g_{n_0}\, y_k \neq x$, and $g_{n_0}\, y_k \longrightarrow x$. This contradicts the fact that x is isolated in C.

If $g_n\, y \neq x$ for all n, then again $g_n\, y \in \overline{C}$ for all n (invariance) and $g_n\, y \longrightarrow x$ yields a contradiction.

Therefore, either C is made out of isolated points only, or no point in C is isolated. In the first instance, $C = \overline{C}$ and C is maximal by Proposition 1 (a).

■

Proposition 3.

If C_1 and C_2 are two invariant control sets and C_1 is maximal, then $C_2 \subset C_1$ or $C_1 \cap C_2 = \phi$. If both C_1 and C_2 are maximal, then $C_1 = C_2$ or $C_1 \cap C_2 = \phi$.

Proof.

Suppose there exist $x \in C_1 \cap C_2$ and $y \in C_2 \setminus C_1$. Then $y \in \overline{Sx} = \overline{C}_2$ and it follows that $\overline{C}_2 = \overline{Sy} \subset \overline{Sx} = \overline{C}_2$, i.e., $\overline{Sy} = \overline{Sx} = \overline{C}_1$ ($\overline{Sy} \subset \overline{Sx}$ by the same argument as in the proof of Proposition 1 (c)). But, by the maximality of C_1, this implies that $y \in C_1$, which is a contradiction.

If both C_1 and C_2 are maximal, it suffices to reverse the roles of C_1 and C_2 in the above reasoning to get $C_1 \cap C_2 = \phi$ or $C_1 \subset C_2$, the latter giving $C_1 = C_2$. ∎

Remark 2.

If C_1 and C_2 are two different nonmaximal invariant control sets, $C_1 \cap C_2$ may not be empty. For example, take $M = \mathbb{S}^1$, the circle in $\mathbb{R}^2$, and identify the elements of $\mathbb{S}^1$ by their polar coordinate θ, $0 \leq \theta < 2\pi$. Then consider the difference equation $\theta_{n+1} = (\theta_n + \varphi) \bmod 2\pi$, where φ is fixed in $(0, \pi)$ and φ is irrational but not a rational multiple of π. It can be shown (see Example 5, p. 98) that, under the semigroup generated by this difference equation, $S\theta$ is dense in $\mathbb{S}^1$ for all $\theta \in \mathbb{S}^1$. Define the invariant control sets C_1 and C_2 by

$$C_1 = \{\theta \in \mathbb{S}^1 : \theta \in \mathbb{Q} \text{ or } \theta = k\varphi, k \in \mathbb{N}, \bmod 2\pi\} \text{ and}$$

$$C_2 = \{\theta \in \mathbb{S}^1 : \theta \in \mathbb{Q} \text{ or } \theta = k\pi, k \in \mathbb{N}, \bmod 2\pi\}.$$

Note that $\overline{C_1} = \overline{C_2} = \mathbb{S}^1$ but that neither C_1 nor C_2 are maximal invariant control sets. Moreover, $C_1 \cap C_2 \neq \phi$ but neither set is included in the other. In this case, $\mathbb{S}^1$ is the unique maximal invariant control set.

Note that, so far, this example is not a controlled difference equation since there are no controls. But a controlled difference equation in $\mathbb{R}^2_0$ exhibiting the same features would be (still using polar coordinates (r, θ), $r > 0$, $\theta \in [0, 2\pi)$,

$$(r_{n+1}, \theta_{n+1}) = (a\, r_n, (\theta_n + \varphi) \bmod 2\pi),$$

with $a > 0$ and φ as above.

Proposition 4.

If C_1 and C_2 are two disjoint invariant control sets (in particular, if C_1 and C_2 are distinct maximal invariant control sets) then

a) if $x \in \overline{C_1} \cap \overline{C_2}$ and C_1 is maximal, then int $\overline{Sx} = \phi$ and

b) int $\overline{C_1} \cap$ int $\overline{C_2} = \phi$ (no maximality requirement).

Proof.

a) Suppose int $\overline{Sx} \neq \phi$. Then Proposition 1 (d) implies $x \in C_1$. But this implies that, under some $g \in \mathcal{S}$, a whole neighborhood of x, U_x, can be mapped into C_1. Since $x \in \partial C_2$, there is some $y \in U_x \cap C_2$ such that $g\,y \in C_1$. This contradicts the invariance of C_1.

b) The result is obvious if int $\overline{C_i} = \phi$ for $i = 1$ or 2.
If int $\overline{C_i} \neq \phi$ for $i = 1, 2$ and int $\overline{C_1} \cap$ int $\overline{C_2} \neq \phi$, then there exists an open set $O \subset$ int $\overline{C_1} \cap$ int $\overline{C_2}$. So, C_1 and C_2 are both dense in O and $x \in C_1 \cap O$ implies $x \in \overline{C_2}$. But, if C_{0_1} and C_{0_2} denote the maximal invariant control sets associated with C_1 and C_2 respectively ($C_i \subset C_{0_i} \subset \overline{C_i}$, see Proposition 1 (e)), we have int $\overline{Sx} =$ int $\overline{C_{0_1}} =$ int $\overline{C_1} \neq \phi$. But then $x \in \overline{C_{0_1}} \cap \overline{C_{0_2}}$ and int $\overline{Sx} \neq \phi$ contradicts part (a) of this proposition. ■

Remark 3.

1) The maximality condition imposed in Proposition 4 (a) is necessary.
To see this, take $M = \mathbb{S}^1$, the unit circle, and $\mathcal{S} = \mathcal{S}(R)$ with the matrix R

defined by $R = \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix}$, φ irrational in $(0, \pi)$ but not rational multiple of π. Identify the elements in $\mathbb{S}^1$ by their angle in radians. Then, for all $\theta \in \mathbb{S}^1$, $S\theta$ is dense in $\mathbb{S}^1$.

Pick $\theta_1 = \pi/6$ and $\theta_2 = \theta_1 + 1$. Both $S\theta_1$ and $S\theta_2$ are nonmaximal invariant control sets in $\mathbb{S}^1$. Moreover, since $\theta_1 - \theta_2 = 1$ and φ is not a rational multiple of π, $S\theta_1 \cap S\theta_2 = \phi$. But $\overline{S\theta_1} \cap \overline{S\theta_2} = \mathbb{S}^1$ and, for all $\theta \in \mathbb{S}^1$, int $\overline{S\theta} = \mathbb{S}^1 \neq \phi$.

For a more detailed discussion of this setup, see Example 5, p. 98.

As in Remark 2, the difference equation on $\mathbb{S}^1$, $s_{n+1} = R\, s_n$, yielding this setup can easily be turned into a controlled difference equation on $\mathbb{R}^2_0$ exhibiting the same behavior: Simply take (still using polar coordinates),

$$(r_{n+1}, \theta_{n+1}) = (a\, r_n, (\theta_n + \varphi) \bmod 2\pi),$$

with $a > 0$.

2) Example 1 shows that, even under maximality, one cannot assert $\overline{C_1} \cap \overline{C_2} = \phi$.

Now, in further discussions, the hypotheses int $C \neq \phi$ or int $\overline{C} \neq \phi$ will become quite important. We therefore collect basic statements giving conditions under which these hypotheses will be satisfied.

Proposition 5.

Let C be an invariant control set. Then

a) int $\overline{Sx} \neq \phi$ for some $x \in \overline{C}$ implies int $\overline{C} \neq \phi$. Conversely, if int $\overline{C} \neq \phi$, then int $\overline{Sx} \neq \phi$ for all $x \in C$.

b) if C is maximal, int $\overline{Sx} \neq \phi$ for all $x \in \partial C$ implies int $C =$ int $\overline{C} \neq \phi$.
In general, the converse is false.

Proof.

a) By Proposition 1 (b), $x \in \overline{C}$ implies $\overline{Sx} \subset \overline{C}$. Hence, int $\overline{Sx} \neq \phi$ implies int $\overline{C} \neq \phi$. Conversely, if $x \in C$, then $\overline{Sx} = \overline{C}$ and so, int $\overline{C} \neq \phi$ implies int $\overline{Sx} \neq \phi$.

b) By Proposition 1 (d), $C = \overline{C}$. Hence $\partial C \subset C$ and the first statement is proved by applying part (a).

Example 1 shows that the converse statement need not hold. ■

The following series of counterexamples shows that the statements made in Proposition 5 cannot be strengthened and further illustrates the relationships between int $\overline{Sx}$, int $\overline{C}$, and int C.

Example 3.

These examples refer to Proposition 5 (a).

1) int $\overline{C} \neq \phi$ does not imply int $\overline{Sx} \neq \phi$ for any $x \in \partial C$ (see Example 1).

2) Even when int $\overline{Sx} \neq \phi$ for all $x \in \overline{C}$, it may be that int $C = \phi$.
This is illustrated by the situation described in Remark 3. Note that, in this case, $C = S\theta_1$ (or $S\theta_2$) is not maximal (see Remark 1).

3) This example shows that, even when C is maximal, int $\overline{Sx} \neq \phi$ for all $x \in C$ does not imply int $C \neq \phi$ (and hence int $\overline{C} \neq \phi$ does not imply int $C \neq \phi$).
Take $M = S^1 \times S^1 = T^2$, the 2-torus. Viewing T^2 as a subset of $\mathbb{C} \times \mathbb{C}$, we can write $x \in T^2$ as $(e^{2\pi i\theta}, e^{2\pi i\gamma})$, $\theta, \gamma \in (0,1)$.
Then consider the diffeomorphisms $\{T_t ; t \geq 0\}$ on T^2 defined by

$$T_t (e^{2\pi i\theta}, e^{2\pi i\gamma}) = \left[e^{2\pi i (\theta+t)}, e^{2\pi i\alpha (\gamma+t)}\right],$$

where α is some fixed irrational number in $[0, 1]$.

Let $\mathcal{S}_1$ be the semigroup $\mathcal{S}(\{T_t ; t \geq 0\})$.
It follows from Boothby (1986, p. 86), that, for all $x \in T^2$, $S_1 x$ is dense in T^2, so that int $\overline{Sx} = T^2 \neq \phi$. Obviously, the negative orbit of any $x \in T^2$,

$$S_1^- x \equiv \{y \in T^2 : T_t\, y = x \text{ for some } t \geq 0\},$$

is also dense in T^2. In fact, $S_1 x$ is the irrational winding line on the torus, starting at x, and is the solution of the differential equation

$$\frac{d\,x(t)}{d\,t} = \begin{bmatrix} \frac{d\,x_1(t)}{d\,t} \\ \\ \frac{d\,x_2(t)}{d\,t} \end{bmatrix} = \begin{bmatrix} 2\pi i\ x_1(t) \\ 2\pi i \alpha\ x_2(t) \end{bmatrix},$$

for $x = x(0) = (x_1(0), x_2(0)) \in T^2 \subset \mathbb{C} \times \mathbb{C}$.
Using the vector field notation of Subsection 2.4, this differential equation can be written as

$$x_* \left[\frac{d}{dt}\right] = X(x(t)) = 2\pi i\, x_1(t) \frac{\partial}{\partial z_1} + 2\pi i \alpha\, x_2(t) \frac{\partial}{\partial z_2}.$$

To obtain the desired example, we change the C^∞ vector field X and use

$$Y = f(x_1(t))\, 2\pi i\, x_1(t) \frac{\partial}{\partial z_1} + f(x_2(t))\, 2\pi i \alpha\, x_2(t) \frac{\partial}{\partial z_2},$$

where $f : \mathbb{C} \longrightarrow \mathbb{C}$ is defined by

$$f(a + bi) = 2ab + i(b^2 - a^2 + 1).$$

Note that, writing $u(a, b) = 2ab$ and $v(a, b) = b^2 - a^2 + 1$,

$$f(a + bi) = u(a, b) + i\, v(a, b))$$

satisfies the Cauchy–Riemann equations and all the partial derivatives of u and v are continuous, which implies that f is analytic and hence that Y is a C^∞ vector field. Moreover, $f(1) = f(-1) = 0$ while, for $a,\ b \in [0,1]$, $a^2 \neq 1$,

$$\operatorname{Re} f(a + bi) > 0 \text{ and}$$
$$\operatorname{Im} f(a + bi) > 0.$$

So, when $(x_1(t), x_2(t)) =$

$$(e^{\pi i}, e^{\pi i}) = (-1,-1), \text{ or}$$
$$(e^{\pi i}, e^{2\pi i}) = (-1,1), \text{ or}$$
$$(e^{2\pi i}, e^{\pi i}) = (1,-1), \text{ or}$$
$$(e^{2\pi i}, e^{2\pi i}) = (1,1),$$

our vector field Y vanishes. The above four points are therefore restpoints of Y. In fact, if we write $\mathcal{S}_2 = \{e^{tY};\ t \geq 0\}$ for the semigroup of diffeomorphisms generated by the vector field Y, we see that points of the form $(e^{k\pi i}, e^{2\pi i\ \gamma})$ or $(e^{2\pi i\ \theta}, e^{k\pi i})$ with $k = 1,\ 2$ and $\theta,\ \gamma \in (0,1]$, satisfy $\operatorname{int} \overline{Sx} = \phi$.

Moreover, if we define the set A by

$$A \equiv \{x \in T^2 : x = (e^{k\pi i}, e^{2\pi i\ \gamma}) \text{ or } x = (e^{2\pi i\ \theta}, e^{k\pi i}),\ k = 1,2 \text{ and } \theta, \gamma \in (0,1]\}$$

and $S_2^- A$ by

$$S_2^- A \equiv \{y \in T^2 : e^{tY}\, y \in A \text{ for some } t \geq 0\},$$

we can see that $y \in S_2^- A$ implies $\operatorname{int} \overline{Sy} = \phi$.

Also, $S_2^- A$ is dense in T^2 since $S_1^- A$ is dense in T^2 and multiplying the components of a vector field by a positive C-valued function does not affect the path of the solution.

Now, for $y \notin \bar{S_2}A$, $\overline{S_2 y} = T^2$ and hence, int $\overline{S_2 y} \neq \phi$. Note that $\left[\bar{S_2}A\right]^c \neq \phi$ since any point of the form $(e^{2\pi i\,\theta}, e^{2\pi i\,\gamma})$ with θ, γ irrational in $(0, 1]$ is in $\bar{S_2}A$.
It is then clear that the maximal invariant control set associated with $\mathcal{S}_2$ is $T^2 \setminus \bar{S_2}A$. Since $\bar{S_2}A$ is dense in T^2, int $C = \phi$ while, for all $x \in C = \left[\bar{S_2}A\right]^c$, int $\overline{S_2 y} = T^2 \neq \phi$.

This example applies to differential equations. In order to obtain a similar result for difference equations, we need to transform $\mathcal{S}_2$, i.e., we need to "discretize" it. Take $\mathcal{S}_3 = \{e^{n\beta Y}\,;\, n \in \mathbb{N}_0\}$ with $\beta \in (0, 1) \setminus \mathbb{Q}$. In other words, instead of considering all times t, we only look at time points which are integer multiples of some fixed irrational number $\beta \in (0, 1)$.
Under the semigroup $\mathcal{S}_3$, the points in T^2 will have their orbits along the irrational winding line, but only in a discrete fashion. Hence the conclusions drawn for the differential equation case will hold provided the density in T^2 of the relevant orbits can be established. We will show that, if $y \in \left[\bar{S_3}A\right]^c$, Sy intersects any open set $U \subset T^2$ (the density of $\bar{S_3}A$ in T^2 is proved similarly).

Suppose there exists an open set $U \subset T^2$ such that $Sy \cap U = \phi$ for some point $y \in \left[\bar{S_3}A\right]^c$. Then, since $\left[\bar{S_3}A\right]^c$ is dense in T^2, there must be a time $t \geq 0$ such that $e^{tY} y \in U$. Write $e^{tY} y = (e^{2\pi i\,\theta_1}, e^{2\pi i\alpha\,\theta_2}) \in U \cap S_2 y$ and pick an open set $V \subset U$ of the form

$$\left\{(e^{2\pi i\gamma_1}, e^{2\pi i\gamma_2}) \in T^2 : \theta_1 - a\beta < \gamma_1 < \theta_1 + a\beta \text{ and } \theta_2 - a\beta < \gamma_2 < \theta_2 + a\beta\,,\, a \leq \tfrac{1}{2}\right\}.$$

Next, pick $n \in \mathbb{N}_0$ such that

$$n\beta < t - a\beta \text{ and}$$
$$(n+1)\,\beta > t + a\beta.$$

If such an n does not exist, we are done since then, $e^{n\beta Y}y \subset V \subset U$. Hence, assuming the existence of such an n, by the axiom of Archimedes, there must be an m such that $m\beta \in (t - a\beta, t + a\beta)$, i.e., $e^{m\beta Y}y \in V \subset U$.

Example 4.

This example refers to Proposition 5 (b).

Consider the setup described in Example 1 (a). In this situation we have two maximal invariant control sets:

$$C_1 = \{(x, y) \in \mathbb{R}_0^2 : x \neq 0, y \neq 0\}, \text{ and}$$

$$C_2 = \{(x, y) \in \mathbb{R}_0^2 : x = 0\} \cup \{(x, y) \in \mathbb{R}_0^2 : y = 0\}.$$

Clearly, in this case, int $C_1 = C_1$ and int $\overline{C_1} = \mathbb{R}_0^2$. This shows that we may have int $C \neq \phi$ and int $\overline{C} \neq \phi$ but int $C \neq$ int $\overline{C}$.

The basic setup of Example 1 also illustrates that int $\overline{Sx} \neq \phi$ for all $x \in \partial C$ is sufficient but not necessary to obtain int $C =$ int $\overline{C} \neq \phi$.

The above results are summarized in Table 1 on the next page.

In view of the previous proposition, it would be natural to expect parallel results concerning int Sx and int C. Unfortunately, statements aiming at the conclusion that int $C \neq \phi$ (or even int $\overline{C} \neq \phi$) based upon int $Sx \neq \phi$ are a lot harder to handle since invariant control sets are defined up to closure. In fact, in most cases, knowing that int $Sx \neq \phi$ will simply give us the same results as knowing that int $\overline{Sx} \neq \phi$ (e.g., int $Sx \neq \phi$ for all $x \in \partial C$ and C maximal imply that int $C \neq \phi$ by Proposition 5 (b)). Nevertheless, one informative result (which requires differentiability) can be given.

Table 3.1. Relationship between int $\overline{Sx} \neq \phi$ and int $C \neq \phi$ [a,b]

int $\overline{Sx} \neq \phi$ for	implies	int $C \neq \phi$	
		when C is not maximal	when C is maximal
some $x \in C$		NO by E 3.2	NO by E 3.3
all $x \in C$		NO by E 3.2	NO by E 3.3
some $x \in \overline{C}$		NO by E 3.2	NO by E 3.3
all $x \in \overline{C}$		NO by E 3.2	YES by P5.b
some $x \in \partial C$		NO by E 3.2	NO by E 3.3
all $x \in \partial C$		NO by E 3.2	YES by P5.b

[a] Note that, whether C is maximal or not, int $\overline{Sx} \neq \phi$ for some $x \in \overline{C}$ implies int $\overline{C} \neq \phi$.

[b] In this table, E 3.2 means Example 3 (2), P5.b means Proposition 5 (b), etc.

Proposition 6.

Let $\mathcal{S}$ be the semigroup generated by the collection $D \equiv \{g_\lambda ; \lambda \in \Lambda\}$ of C^1 maps on M and let C be a maximal invariant control set associated with $\mathcal{S}$. Define $\mathcal{S}^n$ and $S^n x$ by

$$\mathcal{S}^n \equiv \{g_k \circ \cdots \circ g_1 : g_i \in D, 1 \leq i \leq k, k \leq n\} \text{ and}$$

$$S^n x \equiv \{y \in M : y = g_k \circ \cdots \circ g_1 (x) : g_i \in D, 1 \leq i \leq k, k \leq n\}.$$

Assume that there exists $m \in \mathbb{N}$ and an open set $V \subset \overline{C}$ such that, for all open sets $O \subset C$ and all $x \in V \cap C$, $S^m x \cap O \neq \phi$, i.e., that every open set in C can be reached from any point in $V \cap C$ in at most m steps. Also assume that $\|g_\lambda *\| \leq K < \infty$ for all $g_\lambda \in D$. (The norm $\|.\|$ is defined by $\|F_*\| = \sup \{F_* (x) ; x \in M\}$ where F_* is the differential $F_* (x) : T_x (M) \longrightarrow T_{F(x)} (M)$.)

Then int $Sx \neq \phi$ for some $x \in \overline{C}$ implies int $C \neq \phi$.

Proof.

First note that C is necessarily dense in int Sx and that, by Proposition 1 (d), we know that $x \in C$.

We clearly have that int $Sx \cap V \neq \phi$. To prove our result, it is enough to show that $V \cap$ int $Sx \subset C$. We will do this by showing that, for all $y \in V \cap$ int Sx and any open set $U \subset \overline{C}$, $Sy \cap U \neq \phi$. This will imply that Sy is dense in C and, since $y \in \overline{C}$, maximality will then give that $V \cap$ int $Sx \subset C$.

Pick $y \in V \cap$ int Sx. Since C is dense in int Sx, there exists a sequence $\{x_n ; n \geq 1\} \subset C$ such that $x_n \longrightarrow y$ as $n \longrightarrow \infty$. Moreover, by our assumption, to each x_n is associated $g_n \in \mathcal{S}^n$ such that $g_n x_n \in A \subset U$, where A is an open set with $\overline{A} \subset U$. But then $g_n^{-1} (U) \equiv W_n$ is an open set satisfying $x_n \in W_n$ and $Sz \cap U \neq \phi$ for all $z \in W_n$.

If $y \in \bigcup_{n=1}^{\infty} W_n$, we are done.

If $y \notin \bigcup_{n=1}^{\infty} W_n$, then, since we can assume that int $Sx \cap V$ is a bounded set in the metric ρ on M, either the diameter $d(W_n)$ of W_n converges to 0, or $\overline{\lim}_{n\to\infty} d(W_n) > 0$.

In the first case, there exists a sequence $\{z_n, n \geq 1\}$, $z_n \in W_n$, such that

$$\sup \{\rho(z_n, w) ; w \in W_n\} \longrightarrow 0 \text{ as } n \longrightarrow \infty.$$

Now let $g_n z_n = z_n' \in U$ and, for all n, pick $w_n' \in U$ such that $\rho(z_n', w_n') > \epsilon$, $\epsilon > 0$. Since we can assume that U is precompact, Corollary 4.1.2 in Crauel (1987) gives that, for all n,

$$\epsilon < \rho(z_n', w_n') \leq \|g_n *\| \rho(z_n, g_n^{-1} w_n') \leq \|g_n *\| \sup \{\rho(z_n, w) ; w \in W_n\}.$$

Now this supremum converges to zero and the only way this statement can be verified would be for $\|g_n *\|$ to converge to infinity as $n \longrightarrow \infty$. But this is impossible since g_n is the composition of finitely many elements in D, each of them having a linearization in T (M) with bounded norm.

Hence, for some n_0, $y \in W_{n_0}$ and $g_{n_0} y \in U$.

In the second case, i.e., $\overline{\lim}_{n\to\infty} d(W_n) > 0$, we take A such that $\rho(\overline{A}, \partial U) = \delta > 0$ for some δ. By the same argument as above, $\|g_n *\| \leq L < \infty$ for all $n \in \mathbb{N}$, and

$$\rho(g_n x_n, g_n y) \leq \|g_n *\| \rho(x_n, y) \leq L \rho(x_n, y) \longrightarrow 0 \text{ as } n \longrightarrow \infty.$$

But $g_n x_n \in A$, i.e., there is $n_0 \in \mathbb{N}$ with $\rho(g_{n_0} x_{n_0}, g_{n_0} y) < \delta$ and therefore, $g_{n_0} y \in U$.

■

Remark 4.

In Proposition 5 (a) we proved int $C \neq \phi$ under the assumption that C is maximal and int $\overline{Sx} \neq \phi$ for $x \in \partial C$. Proposition 6 requires int $Sx \neq \phi$ only for some $x \in \overline{C}$, but imposes conditions on the number of steps in which open sets can be reached and on the norm of the C^1 maps generating the semigroup $\mathcal{S}$. When $\overline{C}$ is actually compact and $h : M \times U \longrightarrow M$ is C^1 with U open, this norm requirement is satisfied when $\overline{U}$ is compact, and the number of steps to reach any open set is uniformly bounded. Also note that the statement and the proof of this result require the use of paths rather than orbits.

Still related to the above discussion, Example 1 shows it may be that int C is a proper subset of int $\overline{C}$. The next example illustrates that, even when C is maximal, int $C \neq \phi$ does not imply int $Sx \neq \phi$ for any $x \in \overline{C}$.

Example 5.

Let the diffeomorphisms R_n be rotation matrices defined by

$$\begin{bmatrix} \cos n\varphi & \sin n\varphi \\ -\sin n\varphi & \cos n\varphi \end{bmatrix},$$

where φ is a fixed irrational number in $[0, 2\pi)$ not of the form $k\pi/n$, $k, n \in \mathbb{N}_0$. Consider the unit circle S^1 and the semigroup $\mathcal{S}(R_1)$, Then, in fact,

$$\mathcal{S}(R_1) = \{A \in Gl(d, \mathbb{R}) : A = R_n, n \geq 1\}.$$

First note that, for any $\theta \in [0, 2\pi)$ and n, k, and $k' \in \mathbb{N}_0$, $k' > k$,

$$\theta + n\varphi + 2k\pi = \theta + 2k'\pi$$

implies that $\varphi = 2\left(\frac{k'-k}{n}\right)\pi$, which is impossible since, by assumption, no integer

multiple of φ can ever be of the form $k\pi/n$. Hence, for any $\theta \in \mathbb{S}^1$, $S\theta$ cannot be included in any finite subset of the circle and does not retrace its steps.
Suppose there exists an open set $V_0 \subset \mathbb{S}^1$ with $V_0 \cap S\theta = \phi$. We may assume that V_0 is of the form $(\theta - \epsilon,\ \theta + \epsilon)$. Since $V_0 \cap S\theta = \phi$, it must be that, for all $n \geq 1$, the open sets $V_n \equiv (\theta - \epsilon - n\varphi,\ \theta + \epsilon - n\varphi) \pmod{2\pi}$ satisfy $V_n \cap S\theta = \phi$. But, since $\theta - n_1\varphi \neq \theta - n_2\varphi$ for $n_1 \neq n_2$, rotations are isometries, and $\mathbb{S}^1$ is a compact set, $\bigcup \{V_n\ ;\ n \geq 0\} = \mathbb{S}^1$. This implies that $\mathbb{S}^1 \cap S\theta = \phi$ for any $\theta \in \mathbb{S}^1$, which is absurd. Hence, $S\theta$ is actually dense in $\mathbb{S}^1$ and $\overline{S\theta} = \mathbb{S}^1$ for all $\theta \in \mathbb{S}^1$.
$\mathbb{S}^1$ is therefore the maximal invariant control set associated with $\mathcal{S}$. Moreover, $\mathbb{S}^1$ has clearly a nonempty interior (in $\mathbb{S}^1$) while int $S\theta = \phi$ for all $\theta \in \mathbb{S}^1$.

The following double example is concerned with the negative orbit of a point $x \in M$, $S^-x = \{y \in M : g\,y = x \text{ for some } g \in \mathcal{S}\}$. Although we will not make use of S^-x in the subsequent sections, it is interesting to note that, just as in the continuous case (without a Lie algebra condition, which will relate Sx and S^-x), there is no obvious relation between int Sx and int S^-x. In the rest of this section, S^-x will only be used to prove exact controllability results. The first example below shows that one may have int $Sx \neq \phi$ but int $S^-x = \phi$, and the second that one may have int $Sx \neq \phi$, int $S^-x \neq \phi$, and $x \in \overline{\text{int } Sx}$ but nevertheless, $x \notin \overline{\text{int } S^-x}$.

Example 6.

a) Take $M = \mathbb{R}$ and let $\mathcal{S}$ be the semigroup generated by the diffeomorphisms $\{g_{\alpha\beta}\ ;\ \alpha \in (-1, \infty), \beta > 0\}$ where

$$g_{\alpha\beta}(x) = \begin{cases} x+1 & \text{for } x \leq 1 \\ x+1+\alpha \exp\left[-\frac{\beta}{(x-1)^2}\right] & \text{for } 1 < x \leq 2 \\ \vdots & \\ x+1+\alpha \sum_{k=1}^{n} \exp\left[-\frac{\beta}{(x-k)^2}\right] & \text{for } n < x \leq n+1 \end{cases}$$

For any $(\alpha, \beta) \in (-1, \infty) \times \mathbb{R}_0^+$, the map $g_{\alpha\beta}$ is one-to-one and onto. Since, moreover, $g'_{\alpha\beta}$ is positive, $g_{\alpha\beta}$ is a diffeomorphism. Also note that the set $[1, \infty)$ is a maximal invariant control set.

Now, for example, if x is any negative integer, then $[1, \infty) \subset Sx$ and hence int $Sx \neq \phi$. But, clearly, $S^-x = \{y \in \mathbb{R} : y = x - k, k \in \mathbb{N}_0\}$ and int $S^-x = \phi$.

A similar example (but without control set) can be found in Jacubczyk and Sontag (1988, Remark 5.1).

b) Let $M = \mathbb{R}$ and let $\mathcal{S} \subset$ Diff (M) be the semigroup generated by the diffeomorphisms

$$f_{\alpha\beta}(x) = \alpha(e^x + \beta x), \quad \alpha, \beta \in (0, 1).$$

These maps are clearly C^∞ and one–to–one for any fixed $\alpha, \beta \in (0, 1)$, since their derivative is positive. They are also onto and hence, by Proposition 2.1.1, they are diffeomorphisms on $\mathbb{R}$. Now, the positive orbit of 0 is $\mathbb{R}_0^+$ but 0 can only be reached from the points satisfying $e^x = -\beta x$. Such points must clearly be negative and, in fact, they belong to the interval $(-\infty, y)$, where y solves $e^x = -x$ ($x \approx -0.567$). Since such points cannot be reached from any $x \in [y, \infty)$, we have $S^-0 = (-\infty, y)$. Therefore, both the positive and negative orbits of 0 have nonempty interior and $0 \in \overline{\text{int } S0}$ but $0 \notin \overline{\text{int } S^-0}$.

After this investigation of the relationships between the conditions int $\overline{Sx} \neq \phi$, int $Sx \neq \phi$, int $\overline{C} \neq \phi$, and int $C \neq \phi$, we are ready to give a series of results depending on such conditions.

Proposition 7.

Let C be a maximal invariant control set associated with the semigroup $\mathcal{S}$. Then int $C \neq \phi$ implies that C is Borel.

Proof.

We simply reproduce the proof found in Arnold and Kliemann (1987, Remark 3.1.5.) Define the set $D \equiv \{x \in M : Sx \cap \text{int } C \neq \phi\}$ and note that $D \neq \phi$ since int $C \subset \overline{C} = \overline{Sx}$ for all $x \in C$, i.e., $C \subset D$.

First we show that D is open. Indeed, for any $x \in D$, there exists $h_x \in \mathcal{S}$ such that $h_x x \in$ int C. Since h_x is continuous, h_x^{-1} (int C) is an open neighborhood of x which is included in D. This implies that D is open.

Next we prove that $C = D \cap \overline{C}$, i.e., that C is Borel (as it is the intersection of an open and a closed set). Clearly, since $C \subset D$, $C \subset D \cap \overline{C}$.

Moreover, if $x \in D \cap \overline{C}$, then $Sx \cap \text{int } C \neq \phi$. So, there is $y \in Sx \cap$ int C with $\overline{C} = \overline{Sy} \subset \overline{Sx} \subset \overline{C}$, where the first equality holds because $y \in C$ and the second because $x \in \overline{C}$. So, $\overline{Sx} = \overline{C}$ and, by maximality, $x \in C$, i.e., $D \cap \overline{C} \subset C$.

■

Remark 5.

The condition int $C \neq \phi$ in the previous proposition is therefore sufficient to ensure that a maximal invariant control set be Borel. Nevertheless, it is not necessary. For example, on $\mathbb{R}^2_0$, let $\mathcal{S}$ be the semigroup of all rotation matrices. Then, the maximal invariant control sets are all the circles of radius $r > 0$, all of which have empty

interior but are closed sets and hence Borel.

In the later sections, we will concentrate on maximal invariant control sets with nonempty interior and will therefore rely upon Proposition 7 to conclude that these sets are Borel.

Proposition 8.

Let $\mathcal{C}$ be any collection of disjoint invariant control sets (not necessarily maximal) under a semigroup $\mathcal{S}$ acting on a manifold M.

Then int $\overline{C} \neq \phi$ for all $C \in \mathcal{C}$ implies that $\mathcal{C}$ is a countable set. In particular, if all the maximal invariant control sets under $\mathcal{S}$ have nonempty interior, then there are at most countably many such sets in M.

Proof.

By Proposition 4 (b), if C_1 and C_2 are different invariant control sets, we have int $\overline{C_1} \cap$ int $\overline{C_2} = \phi$. Hence, it is enough to recall that, by definition, M has a countable basis of open sets to see that the number of maximal invariant control sets is at most countably infinite. For maximal invariant control sets, the statement follows from the fact that such sets must be disjoint or they are equal (see Proposition 3).

■

Example 7.

This example gives a nonlinear situation with uncountably many maximal invariant control sets even though one of them has a nonempty interior.

Let $M = \mathbb{R}$ and let $\mathcal{S}$ be the semigroup generated by the collection of diffeomorphisms $\{g_\alpha ; \alpha \in (-1, \infty)\}$.

$$g_\alpha x = \begin{cases} x & \text{for } x \leq 1 \\ x + \alpha \exp\left[-\frac{1}{(x-1)^2}\right] & \text{for } x > 1, \end{cases}$$

Then, for $x \leq 1$, each singleton $\{x\}$ is a maximal invariant control set while the open set $(1, \infty)$ is another maximal invariant control set with nonempty interior.

Proposition 9.

Let C be an invariant control set (not necessarily maximal) under the semigroup $\mathcal{S}$ and assume that int $C \neq \phi$. Then the following holds:

a) int C is C-invariant provided that, for all $y \in \partial C$ and all open neighborhoods U of y, there exists $z \in U \cap (\overline{C})^c \neq \phi$. Moreover, int $\overline{C}$ is always C-invariant.

b) $\overline{\text{int } C} = \overline{C}$.

Recall that $\overline{C}$ is C–invariant (Proposition 1 (b)) but that neither C nor int C need to be C–invariant (Example 2).

Proof.

a) Suppose that int C is not C-invariant, i.e., that there exists $g \in \mathcal{S}$ with $g\,y \notin$ int C for some $y \in$ int C. Since $\overline{Sy} = \overline{C}$, this clearly means that $g\,y \in \overline{C} \setminus$ int $C = \partial C$. Then, every neighborhood U of $g\,y$ contains points in $(\overline{C})^c \neq \phi$. This implies that there is a point z in every neighborhood V of y for which $g\,z \in (\overline{C})^c$, which contradicts the fact that C is an invariant control set.
That int $\overline{C}$ is C–invariant follows immediately from the above argument (with $\overline{C}$ replacing C). Indeed, when $C = \overline{C}$, there always exists $z \in U \cap (\overline{C})^c \neq \phi$ for all $y \in \partial C$ and all open neighborhoods U of y.

b) Clearly $\overline{\text{int } C} \subset \overline{C}$. So we need only show that $\overline{C} \subset \overline{\text{int } C}$. Let x be any point in $\overline{C}$. If $x \in$ int C, we are done.

If $x \in \partial C$, Sy dense in $\overline{C}$ for $y \in \text{int } C$ implies there is $\{g_n ; n \geq 1\} \subset \mathcal{S}$ and $z \in \text{int } C$ such that $g_n z \longrightarrow x$ as $n \longrightarrow \infty$. But, by the proof of part (a) just above, for all $n \geq 1$, $g_n(z) \in \overline{\text{int } C}$ and hence, $x \in \overline{\text{int } C}$.

■

Remark 6.

Whether C is maximal or not, replacing C by $\overline{C}$ (even if $\overline{C}$ itself is not a control set) in the proof of Proposition 9 (b) shows that int $\overline{C} \neq \phi$ implies $\overline{\text{int } \overline{C}} = \overline{C}$.

Statements on control sets like int $C \neq \phi$ implies $\overline{\text{int } C} = \overline{C}$ (or int $\overline{C} \neq \phi$ implies $\overline{\text{int } \overline{C}} = \overline{C}$) do not have their equivalent for orbits without further hypotheses. This is obvious from the example below. Nevertheless, the following proposition shows that, under some additional assumptions, one can prove that $\overline{\text{int } Sx} = \overline{Sx}$.

Example 8.

Take $M = \mathbb{R}_0^+ \times \mathbb{R}_0^+$ and consider the controlled difference equation

$$x_{n+1} = \begin{bmatrix} \alpha + \beta f(\alpha) & 0 \\ 0 & \alpha + \gamma f(\alpha) \end{bmatrix} x_n,$$

with $\alpha > 1$, γ, $\beta > 0$, and where

$$f(y) = \begin{cases} 0 & y \leq 2 \\ \exp\left[-[y-2]^{-2}\right] & y > 2 \end{cases}.$$

Consider the point $x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

We can then see that $\overline{Sx} = \{(x, y) \in M : x \geq 2,\ y \geq 2\} \cup \{(x, y) \in M : 1 \leq x = y\}$,

int Sx = int $\overline{Sx}$ = $\{(x, y) \in M : x > 2,\ y > 2\}$, and $\overline{\text{int } Sx} = \{(x, y) \in M : x \geq 2,\ y \geq 2\}$.

So, $\overline{\text{int } Sx} \neq \overline{Sx} \neq \overline{\text{int } \overline{Sx}}$.

Also see Example 5.

Proposition 10.

Let C be a maximal invariant control set under the action of the semigroup $\mathcal{S}$ on a manifold M. Assume that, for all $x \in \overline{C}$, int $Sx \neq \phi$ and that either the elements of $\mathcal{S}$ are open maps (e.g., homeomorphisms) or that int $C \cap$ int $S^{-}x \neq \phi$ for all $x \in C$. Then $\overline{\text{int } Sx} = \overline{Sx}$ for all $x \in \overline{C}$.

Proof.

First remember that int $\overline{Sx} \neq \phi$ for all $x \in \overline{C}$ implies $\overline{Sx} = \overline{C} = C$ (see Proposition 1). Hence we know that $\overline{\text{int } Sx} \subset \overline{Sx} = C$, and all we have to show is that $C \subset \overline{\text{int } Sx}$. By the above, int $C \neq \phi$ and, by Proposition 9, $\overline{\text{int } C} = C$. We thus have to show that, for all open sets $U \subset$ int C, $U \cap$ int $Sx \neq \phi$.

Choose such a set U and any $x \in C = \overline{C}$. Note that C is dense in U. Furthermore, int $Sx \subset C \neq \phi$, and thus, for any point $u \in U \subset C$, there exist $y \in$ int Sx and a sequence $\{g_n ; n \in \mathbb{N}\} \subset \mathcal{S}$ such that $g_n y \longrightarrow u$. This means that, for $n \geq K$, $g_n y \in U$. If $g_n y \in$ int Sx, we are done. But, if $\mathcal{S}$ is a semigroup of open maps, this is obvious because open mappings will map interior points (of Sx) into interior points (of Sx). If the maps in $\mathcal{S}$ are not necessarily open but int $S^{-}u \cap$ int $C \neq \phi$ for all $u \in U$, then, since Sx is dense in C, $Sx \cap$ int $S^{-}u$ for all $u \in U$ and therefore, we immediately have that $U \subset$ int Sx.

■

Corollary 1.

If, for all $x \in \overline{C}$, int $Sx \neq \phi$ and either the elements of $\mathcal{S}$ are open maps (e.g., homeomorphisms) or int $C \cap$ int $S^{-}x \neq \phi$, then $Sx =$ int C for all $x \in$ int C.

Proof.

Since int $Sx \neq \phi$ for all $x \in \overline{C}$, $\phi \neq$ int $C =$ int $\overline{C}$ is C-invariant by Proposition 9 (a). Also, the argument held in the proof of Proposition 10 implies that, for all $x \in \overline{C} = C$, int $C =$ int Sx. For $x \in$ int C, the result then simply follows from the inclusions $Sx \subset$ int $C \subset$ int $Sx \subset Sx$. ∎

Proposition 11.

If int $\overline{C} \neq \phi$ for some invariant control set under a semigroup $\mathcal{S}$ acting on M, then:

a) int $\overline{C} = \bigcup_{i=1}^{\infty} O_i$ where the O_i's are disjoint, open, and connected sets,

b) $\overline{C} = \overline{\bigcup_{i=1}^{\infty} \overline{O_i}}$,

c) $x \in \overline{O_i} \cap \overline{O_j}$ for some (i, j), $i \neq j$, implies $x \in \partial C$, and

d) if, for two components O_i and O_j, there is $x \in \overline{O_i} \cap \overline{O_j}$ with int $\overline{Sx} \neq \phi$, then $i = j$ or, for each O_k, there exists $g_k \in \mathcal{S}$ such that $g_k (O_i \cup O_j) \subset O_k$. In fact, if $g x \in O_k$ for some $g \in \mathcal{S}$, then $g (O_i \cup O_j) \subset O_k$.

Proof.

a) In a locally compact space with countable basis, every open set is the countable union of its connected open components.

b) By Remark 6, $\overline{\text{int } \overline{C}} = \overline{C}$ and hence, $\overline{C} = \overline{\bigcup_{i=1}^{\infty} O_i} = \overline{\bigcup_{i=1}^{\infty} \overline{O_i}}$.

c) $x \in \overline{O_i} \cap \overline{O_j}$ implies $x \in \partial \overline{C} = \partial C$ or $x \in \text{int } \overline{C}$.

But $x \in \text{int } \overline{C}$ implies that $x \in \bigcup_{i=1}^{\infty} O_i$ and, since the O_i's are disjoint, this is impossible.

d) Take $x \in \overline{O_i} \cap \overline{O_j}$ with $\text{int } \overline{Sx} \neq \phi$. Then, by (c), $i = j$ or $x \in \partial C$. Since $\text{int } \overline{Sx} \neq \phi$, Proposition 1 (d) implies that $x \in C_0$ where C_0 denotes the unique maximal invariant control set corresponding to C. We then have that $C \subset C_0 \subset \overline{C}$ (see Proposition 1 (e)) and hence $\text{int } \overline{C_0} = \text{int } \overline{C} = \bigcup_{i=1}^{\infty} O_i$. Therefore, $\overline{Sx} = \overline{C_0}$ implies that Sx is dense in $\bigcup_{i=1}^{\infty} O_i$. It follows that, for each k, there must exist $g_k \in \mathcal{S}$ such that $g_k x \in O_k$. Since g_k is continuous, it maps connected sets into connected sets. This as well as the C-invariance of $\text{int } \overline{C}$ (see Proposition 9 (a)) imply that the connected set $O_i \cup O_j \cup \{x\}$ must be mapped into O_k and hence, the result follows. ■

Remark 7.

The statement of Proposition 11 (c) says that, if $\overline{O_i} \cap \overline{O_j} \neq \phi$, the set $\partial O_i \cap \partial O_j$ must be "sparse enough" in the sense that any point in $\partial O_i \cap \partial O_j$ must contain points of $(\overline{C})^c$ in any of its neighborhoods. This means that

1) in $\mathbb{R}^1$, $\overline{O_i} \cap \overline{O_j} = \phi$ for all $i \neq j$,

2) in $\mathbb{R}^2$, $\overline{O_i} \cap \overline{O_j}$, $i \neq j$, cannot contain a line segment,

etc.

Working on a compact manifold will enable us to prove some stronger results than the ones previously given. This is the purpose of the following propositions:

Proposition 12.

Given a semigroup $\mathcal{S}$ acting on some compact manifold M, then

a) for all $x \in M$, there exists a maximal invariant control set C_x satisfying

$$C_x = \overline{C_x} \subset \overline{Sx}, \text{ and}$$

b) $C = \bigcap \{\overline{Sx} ; x \in M\}$ is a maximal invariant control set, provided $C \neq \phi$.

Proof.

This result is a basically trivial enlargement of Lemma 3.1 in Arnold et al. (1986a). Note that the Condition (B) stated in this reference for this result (and which is equivalent to the transitive action of $\mathcal{G}(\mathcal{S})$ on M (see Subsection 3.3) by Proposition 2.1 in this same paper) is not necessary. Let us also mention that compactness is only required for part (a) of this proposition.

■

Corollary 2.

Let $\mathcal{S}$ be a semigroup acting on some compact manifold M.

a) If $C = \bigcap \{\overline{Sx} ; x \in M\} \neq \phi$ and int $C \neq \phi$, then C is the unique maximal invariant control set associated with the action of $\mathcal{S}$.

b) Conversely, if a unique maximal invariant control set C exists, then C must be of the form $\bigcap \{\overline{Sx} ; x \in M\} \neq \phi$.

Proof.

a) If C' is another maximal invariant control set, $C \cap C' = \phi$ by Proposition 3, while $x \in C'$ implies $\phi \neq \overline{Sx} \cap C = \overline{C'} \cap \overline{C}$. It follows, by Proposition 4 (a), that $\operatorname{int} \overline{Sx} = \phi$. But then $\operatorname{int} C \subset \operatorname{int} \overline{Sx} = \phi$ contradicts $\operatorname{int} C \neq \phi$ and hence C must be the unique maximal invariant control set.

b) If there exists $x \in M$ with $\overline{Sx} \cap C = \phi$, then, by Proposition 12, there exists another maximal invariant control set $C' \subset \overline{Sx}$ and, by Proposition 3, $C \cap C' = \phi$. Since this contradicts the uniqueness of C, it must be that $\overline{Sx} \cap C \neq \phi$ for all $x \in M$. Moreover, for $y \in \overline{Sx} \cap C$, $\overline{Sy} = \overline{C}$ implies that $\overline{C} \subset \overline{Sx}$ (since $\overline{Sy} \subset \overline{Sx}$) and hence $\overline{C} \subset \bigcap \{\overline{Sx} ; x \in M\} \neq \phi$.

Conversely, if $w \in \bigcap \{\overline{Sx} ; x \in M\}$, then $w \in \overline{Sx}$ for all $x \in M$ and, in particular, for $x \in C$. Therefore $w \in \overline{C}$ (otherwise C cannot be an invariant set) and $\overline{C} \supset \bigcap \{\overline{Sx} ; x \in M\}$

That $C = \overline{C}$ follows immediately from $\overline{C} = \bigcap \{\overline{Sx} ; x \in M\} \subset \overline{Sz} \subset \overline{C}$ for all $z \in \partial C$ (the second inclusion resulting from the C–invariance of $\overline{C}$ (see Proposition 1 (b)), which, by maximality, implies that $z \in C$ for all $z \in \partial C$. ■

Remark 8.

The compactness assumption made in the Proposition 12 is crucial. Indeed, if $M = \mathbb{R}$ and $\mathcal{S}$ is generated by the collection of diffeomorphisms $\{g_a ; a > 1\}$, where $g_a x = x + a$, there cannot be any invariant control set in $\mathbb{R}$ since, for any $x \in \mathbb{R}$ and any $y \in Sx$, $|x - y| > 1$.

Remark 9.

The projection onto $\mathbb{S}^1$ of Example 1 (a) shows that the assumption int $C \neq \phi$ in Corollary 2 (a) cannot be dropped. Indeed, identifying the elements of $\mathbb{S}^1$ with their angle in radians, $C = \bigcap \{\overline{S\theta}\,;\, \theta \in \mathbb{S}^1\} = \{0, \pi/2, \pi, 3\pi/2\}$ is one maximal invariant control set, while $\mathbb{S}^1 \setminus C$ is another maximal invariant control set.

Proposition 13.

Let $\mathcal{S}$ be a semigroup acting on the compact manifold M. Then int $\overline{C} \neq \phi$ for all maximal invariant control sets implies that all the maximal invariant control sets are closed and therefore satisfy int $C \neq \phi$.

Proof.

By Proposition 1 (d), it is enough to show that, for all $y \in \partial C$, int $\overline{Sy} \neq \phi$. Indeed, this will imply that $y \in C$, i.e., that $\partial C \subset C$.

If int $\overline{Sy} = \phi$, applying Proposition 12 (a) yields that there must exist a maximal invariant control set C_y satisfying $C_y = \overline{C_y} \subset \overline{Sy}$. But this implies that int $\overline{C_y} = \phi$, which is a contradiction.

■

Remark 10.

Again, the statement of the above proposition does not hold if M is not compact. To see this, take $M = \mathbb{R}^2_0$ and let $\mathcal{S}$ be generated by the collections of matrices $\{A_{ab\lambda}\,;\, a \neq 0, b \in \mathbb{R}, \text{and } \lambda \geq 1\}$, where $A_{ab\lambda} = \begin{bmatrix} a & 0 \\ b & \lambda \end{bmatrix}$.

Under this setup, the vertical axis $V \equiv \{(x, y) \in \mathbb{R}^2_0 : x = 0\}$, is an eigenspace for all the matrices generating $\mathcal{S}$ and hence, for all the elements of $\mathcal{S}$. It follows that V is invariant under $\mathcal{S}$. But, since, for all $z \in V$, $|A_{ab\lambda}\, z| = |\lambda\, z| \geq |z|$, V is certainly

not an invariant control set.

On the other hand, for all $z_1, z_2 \notin V$, there exists a matrix $A_{ab\lambda} \in \mathcal{S}$ such that $A_{ab\lambda} z_1 = z_2$ and hence, $\mathbb{R}_0^2 \setminus V$ is the unique maximal invariant control set under $\mathcal{S}$. We then see that $C \equiv \mathbb{R}_0^2 \setminus V$ satisfies int $\overline{C} \neq \phi$ but that C is not closed.

Before leaving this section, let us make a few comments concerning the results obtained here and their counterparts in the continuous time case (see, e.g., Arnold et al. (1986a, Section 3), Arnold and Kliemann (1987, Section 3), as well as Kliemann (1987, Section 2)). In most cases, the results obtained above correspond to or extend similar results for the continuous case. But there are differences.

In the continuous case, orbits are always path connected and maximal invariant control sets are path connected under some Lie algebra condition which guarantees that int $Sx \neq \phi$ (and more) for all $x \in M$ (Kliemann (1987, Lemma 2.1)). This is not necessarily the case in our discrete setup where orbits may not be path connected (see Example 6 (a)) and where maximal invariant control sets can be disconnected even when int $Sx \neq \phi$ for all $x \in M$. The following example illustrates such a situation:

Example 9.

Consider the controlled dynamical system on $\mathbb{R}_0^2$

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_n \\ y_n \end{bmatrix} \quad \text{with } \alpha, \beta < 0.$$

Then the set $C = \{(x, y) \in \mathbb{R}_0^2 : x \cdot y \geq 0\}$, i.e., the union of the upper right and lower left closed quadrants, is a disconnected maximal invariant control set. Note that the

role of the matrix $\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$, which has both its distinct eigenspaces included in int C, is to ensure that the coordinates axes (without the origin) are included in C, and hence are not maximal invariant control sets (compare with several of the situations described in Example 1).

In fact, working on $M = O \subset \mathbb{R}^d$ with O open, Meyn (1989, Proposition 2.2) states that, under the assumptions that int $Sx \neq \phi$ for all $x \in M$ and that the map describing the dynamics of the system, $h : O \times U \longrightarrow O$, is C^∞, any maximal invariant control set C is of the form $C = \bigcup_{i=1}^{m} C_i$ with the C_i's closed and disjoint. Moreover, for $g \in \mathcal{S}$ (the semigroup generating C), $g(C_i) = C_{i+1} \pmod m$, i.e., we have a periodicity in Meyn's terminology. Moreover, in his Corollary 2.2.b, Meyn states that, if U is connected, maximal invariant control sets are connected if and only if they are aperiodic, i.e., if and only if $m = 1$.

Remark 11.

Note that this last result from Meyn does not hold if the assumption int $Sx \neq \phi$ for all $x \in M$ is dropped or if U is disconnected. Indeed, if int $Sx = \phi$ for some $x \in M$, C may not be a finite union of closed sets (see, e.g., Example 1 (a) and Example 2 (a)). Note that the disconnected components of the control sets in Example 1 (a) still exhibit periodicity while this is not the case in Example 2 (a) (which does not exclusively use diffeomorphisms ; we were not able to construct a similar example for which all the $h(.,u)$ maps, $u \in U$, are diffeomorphisms). The controlled dynamical system described in Example 9 but with $\alpha, \beta \neq 0$ gives a counterexample for the case where U is disconnected.

Finally, note that Meyn's result implies that, under his assumptions, a periodic behavior cannot take place if the identity map belongs to $\mathcal{S}$.

In the continuous case, the Lie algebra condition imposed on the dynamics of the systems yields both int $Sx \neq \phi$ and int $S^{-}x \neq \phi$ for all $x \in M$ (see Isidori (1985, Corollary 2.13, p.66)). Moreover, by allowing the dynamics of a continuous time control system to act for arbitrarily small times, it is clear that we necessarily have $x \in \overline{\text{int } Sx} \cap \overline{\text{int } S^{-}x}$. In the discrete time case, the situation is different: There are two separate Lie algebra conditions, one for the positive and one for the negative orbits of the system (see Subsection 3.3). The two Lie algebra conditions are not equivalent. The Lie algebra condition related to the positive orbits of the system will ensure that int $Sx \neq \phi$ for all $x \in M$ but does not guarantee that int $S^{-}x \neq \phi$ for all $x \in M$ (see Example 6 (a)). Moreover, Examples 6 (a) and 6 (b) show that we may have $x \notin \overline{\text{int } Sx}$ and $\overline{\text{int } Sx} \cap \overline{\text{int } S^{-}x} = \phi$, respectively. Note that, using the same setup as in Example 6 (b) but with the semigroup $\mathcal{S}$ generated by the inverse maps, we have both int $Sx \neq \phi$ and int $S^{-}x \neq \phi$ for all $x \in \mathbb{R}$, but still $\overline{\text{int } S0} \cap \overline{\text{int } S^{-}0} = \phi$ and $0 \notin \overline{\text{int } S0}$. Hence, in the discrete time case, even under both the "forward" and "backward" Lie algebra conditions, such "disconnected" behaviors can still be seen.

3.2. Semigroups of Invertible Matrices Acting on $\mathbb{R}_0^d$

In the previous section, we usually attempted to construct examples and counterexamples involving the action of subsemigroups from Gl $(d, \mathbb{R})$ on $\mathbb{R}_0^d$ (or $\mathbb{S}^{d-1}$ and $\mathbb{P}^{d-1}$), $d \geq 2$, with the understanding that these subsemigroups could arise from

difference equations. These are indeed the easiest situations to describe and understand. Nevertheless, in many cases, such nice examples were not found and it seems reasonable to hypothesize that many of the results given in Subsection 3.1 could be strengthened when limited to the linear action of Gl $(d, \mathbb{R})$ on $\mathbb{R}_0^d$ $(d \geq 2)$.

Investigations of this type are somewhat outside the main scope of this work and were therefore not carried very far. Only a couple of results related to this problem were obtained, which may be useful for an eventual future in depth study in this area.

Proposition 1.

Let S be a subsemigroup of Gl $(d, \mathbb{R})$ acting on $\mathbb{R}_0^d$ $(d \geq 2)$ and assume that C is a closed maximal invariant control set under S satisfying int $C \neq \phi$.

Then C is of the form $\overline{\bigcup_{i=1}^{\infty} O_i}$, where $\{O_i ; i \in \mathbb{N}_0\}$ is a collection of connected and disjoint open sets satisfying

if $d = 2$, then $\overline{O_i} \cap \overline{O_j} = \phi$ for $i \neq j$, or

if $d > 2$, then $z \in \overline{O_i} \cap \overline{O_j} \neq \phi$ for $i \neq j$ implies that

$$L_z^+ \equiv \{x \in \mathbb{R}_0^d : x = \lambda z,\ \lambda > 0\} \subset \overline{O_i} \cap \overline{O_j} \neq \phi.$$

Proof.

The main part of the above statement follows directly from Proposition 3.1.11 (b). Suppose $\overline{O_{i_0}} \cap \overline{O_{j_0}} \neq \phi$ for some (i_0, j_0) and let $z \in \overline{O_{i_0}} \cap \overline{O_{j_0}}$. We will show that either

$L_z^+ \subset \overline{O_{i_0}} \cap \overline{O_{j_0}}$ and $d > 2$ or

such a z does not exist and d = 2.

Since $z \in \partial C \subset C$ and $\overline{Sy} = \overline{C} = C$ for all $y \in C$, there exists $\{B_n ; n \geq 1\} \subset \mathcal{S}$ such that $B_n z$ belongs to any given $\frac{1}{n}$-neighborhood of z, i.e., $|B_n z - z| < \frac{1}{n}$.

Define $L_z \equiv \{x \in \mathbb{R}_0^d : x = \lambda z,\ \lambda \in \mathbb{R}_0\}$.

Claim : $L_z \cap (O_{i_0} \cap O_{j_0}) = \phi$.

Suppose L_z is such that $L_z \cap O_{i_0} \neq \phi$ or $L_z \cap O_{j_0} \neq \phi$.
We will use O_{i_0} and show this leads to a contradiction. The reasoning for O_{j_0} is identical. Note that $L_z \cap O_{i_0} \neq \phi$ implies that the intersection is of the form $(\frac{1}{\alpha} z, z)$ or $(z, \alpha z)$ where $\alpha > 1$. (Note that z may not belong to $L_z \cap O_{i_0}$ since $z \notin O_{i_0}$.)
Let $y \in L_z \cap O_{i_0} \neq \phi$ be such that $y = \lambda z$ for some $|\lambda| \in (0, 2]$ and let V_δ be a δ-neighborhood of y such that $V_\delta \subset O_{i_0}$. Then, for $\frac{1}{n} < \frac{\delta}{2}$,

$$|B_n y - y| \leq |\lambda|\ |B_n z - z| \leq 2 \frac{1}{n} < \delta,$$

so that $B_n y \in V_\delta \subset O_{i_0}$.

But, since $z \in \overline{O_{j_0}}$ and $\overline{Sz} = \overline{C}$ with $\bigcup_{i=1}^{\infty} \overline{O_i} \subset \overline{\bigcup_{i=1}^{\infty} \overline{O_i}} = \overline{C}$, we can pick the sequence $\{B_n\}$ such that $B_n z \in O_{j_0}$ for all n. But, since O_{i_0} and O_{j_0} are disjoint, this yields a contradiction because the connected segment [y, z] should be mapped, together with z, into the path connected set O_{j_0}. (End of Claim.)

Now there is $A \in \mathcal{S}$ such that $A z \in O_{i_0}$.
Let $L_{Az}^+ \equiv \{x \in \mathbb{R}_0^d : x = \lambda A z,\ \lambda > 0\}$ and define the closed set K^+ by

$$K^+ \equiv L_{Az}^+ \cap \overline{O_{i_0}} \cap \{x \in \mathbb{R}_0^d : |x| \geq |A z|\}.$$

Again we know there must be a sequence $\{D_n ; n \in \mathbb{N}_0\} \subset \mathcal{S}$ such that $|D_n A z - z| < \frac{1}{n}$. Moreover, since, fo each n, D_n is a linear map, D_n must map K^+ into $L^+_{D_n Az}$. Define $M^+_n \equiv D_n K^+ \subset L^+_{D_n Az}$ and pick $x \in K^+$, i.e., $x = \lambda A z$ for some $\lambda > 0$.

Claim: $\lim_{n\to\infty} |D_n x - z| > 0$.

$$\text{If } 0 = \lim_{n\to\infty} |D_n x - z| = \left|\lambda \lim_{n\to\infty} D_n A z - \lim_{n\to\infty} D_n A z\right|$$

$$\geq |\lambda| \left|\lim_{n\to\infty} D_n A z\right| - \left|\lim_{n\to\infty} D_n A z\right|,$$

we have $\left|\lim_{n\to\infty} D_n A z\right| = 0$. Under the assumption that $A z = \sum_{i=1}^{d} \alpha_i e_i$ with $|\alpha_i| > 0$ for $i = 1, \ldots, d$ (which we can always arrange since $S(A z)$ is dense in int $C \neq \phi$), this further implies that $|z| = 0$, i.e., that $z = 0$. This is impossible since $z \in \mathbb{R}^d_0$. (End of Claim.)

Since $\lim_{n\to\infty} |D_n x - z| > 0$ for $x \in K^+$, we conclude that, if $z \in \overline{O_{i_0}} \cap \overline{O_{j_0}}$, the entire line segment $M^+ \equiv \lim_{n\to\infty} M^+_n \subset \lim_{n\to\infty} L^+_{D_n Az} = L^+_z$ must be included in $\overline{O_{i_0}} \cap \overline{O_{j_0}}$. Repeating the argument for $K^- \equiv L^+_{Az} \cap \overline{O_{i_0}} \cap \{x \in \mathbb{R}^d_0 : |x| \leq |A z|\}$, shows that in fact, for some $\alpha > 1 > \beta > 0$, the closed interval

$$L^z_{\alpha\beta} \equiv \{x \in \mathbb{R}^d_0 : x = \lambda z,\ \alpha \geq \lambda \geq \beta\} \subset \overline{O_{i_0}} \cap \overline{O_{j_0}}.$$

Claim: $L^+_z \subset \overline{O_{i_0}} \cap \overline{O_{j_0}}$.

By simply repeating the above reasoning for $z' = \alpha z$ or $z' = \beta z$, it is clear that $L^z_{\alpha_n \beta_n} \in \overline{O_{i_0}} \cap \overline{O_{j_0}}$ for a sequence $\{(\alpha_n, \beta_n) ; n \in \mathbb{N}_0\}$, $\alpha_n > 1 > \beta_n > 0$. We will show that $\alpha_n \longrightarrow \infty$ (that $\beta_n \longrightarrow 0$ follows from a similar argument).

Suppose that $\alpha_n \rightarrow a < \infty$, i.e., $\left[\overline{\bigcup_{n=1}^{\infty} L^z_{\alpha_n \beta}}\right]^c \cap L^z_{\infty \beta} = (a\,z, \infty) \subset \left[\overline{O_{i_0}} \cap \overline{O_{j_0}}\right]^c$. Then $a\,z \in \overline{O_{i_0}} \cap \overline{O_{j_0}}$. But this implies that the argument above can be applied to the point $(a\,z)$ instead of z to yield that

$$L^{az}_{\alpha\beta} \equiv \{x \in \mathbb{R}^d_0 : x = \lambda\, a\, z,\ \alpha \geq \lambda \geq \beta\} \subset \overline{O_{i_0}} \cap \overline{O_{j_0}}$$

for some $\alpha > 1 > \beta > 0$, i.e., $a\,z \in \overline{O_{i_0}} \cap \overline{O_{j_0}}$. The same argument for β then means that $L^+_z = L^z_{\infty\, 0} \subset \overline{O_{i_0}} \cap \overline{O_{j_0}}$. (End of claim.)

This last claims basically ends the proof. Indeed, for $d = 2$, Proposition 3.1.11 and Remark 3.1.7 imply that $\overline{O_{i_0}} \cap \overline{O_{j_0}}$ must be empty for $i \neq j$. ∎

Remark 1.

For the case $d = 1$ (unconnected state space), the result $\overline{O_i} \cap \overline{O_j} = \phi$ for $i \neq j$ is still valid by Proposition 3.1.11 and Remark 3.1.7.

Proposition 2.

Let $\mathcal{S}$ be a semigroup from Gl $(d, \mathbb{R})$ acting on $\mathbb{R}^d_0$ $(d \geq 2)$ and let C be a maximal invariant control set under $\mathcal{S}$ containing some interval $[x, \lambda\, x]$, $\lambda > 0$.

Then $z \in C$ implies that $L^+_z \equiv \{y \in \mathbb{R}^d_0 : y = \lambda\, z,\ \lambda > 0\} \subset C$.

Proof.

First we show that $z \in C$ implies there exists $\alpha > \beta > 0$ such that

$$L^z_{\alpha\beta} \equiv \{y \in \mathbb{R}^d_0 : y = \lambda\, z,\ \alpha \geq \lambda \geq \beta\} \subset C.$$

To see this, pick $v \in (x, \lambda\, x)$. Then there exists a sequence $\{B_n\ ;\ n \in \mathbb{N}_0\} \subset \mathcal{S}$ such

that $\lim_{n\to\infty} B_n v = z$. Also, using the same argument as in Proposition 1, $\lim_{n\to\infty} B_n [x, \lambda x]$ must be a closed interval of the form $L^z_{\alpha\beta} \subset \overline{C}$.

Hence some dense subset of C is included in $L^z_{\alpha\beta}$.

Suppose there is $z_0 \in L^z_{\alpha\beta} \cap C^c$. Since $z_0 \in L^z_{\alpha\beta}$, there is $x_0 \in [x, \lambda x]$ with

$$\lim_{n\to\infty} B_n x_0 = z_0, \text{ i.e., } z_0 \in \overline{Sx_0}.$$

Next, there is $\{z_0^n ; n \in \mathbb{N}_0\} \subset C \cap L^z_{\alpha\beta}$ with $z_0 = \lim_{n\to\infty} z_0^n$ and, for each n, $\overline{Sz_0^n} = \overline{C}$.

Moreover we can write $k_n z_0 = z_0^n$ for some $k_n \in \mathbb{R}_0$ with $k_n \to 1$ as $n \to \infty$. So, since $\mathcal{S} \subset Gl(d, \mathbb{R})$, $\overline{C} = \lim_{n\to\infty} \overline{Sz_0^n} = \lim_{n\to\infty} k_n \overline{Sz_0} = \overline{Sz_0}$ as well, which, by maximality, gives that $z_0 \in C$. Since this contradicts $z_0 \in L^z_{\alpha\beta} \cap C^c$, we conclude that $L^z_{\alpha\beta} \cap C^c = \phi$, i.e., $L^z_{\alpha\beta} \subset C$.

That $L^+_z \subset C$ then follows from the same argument as at the end of the proof of Proposition 1.

■

<u>Corollary 1.</u>

If $\mathcal{S}$ is a semigroup from $Gl(d, \mathbb{R})$ acting on $\mathbb{R}^d_0$ $(d \geq 2)$ and C is a maximal invariant control set under $\mathcal{S}$ such that int $C \neq \phi$, then

$$\text{int } C = \bigcup_{i=1}^{\infty} O_i = \bigcup_{i=1}^{\infty} \bigcup_{z\in O_i}^{\infty} L^+_z,$$

where the O_i's are the open, disjoint, and connected components of int C.

Proof.

int $C = \bigcup_{i=1}^{\infty} O_i$ comes from Proposition 3.1.11 (a) and Proposition 2 above shows that

$$O_i = \bigcup_{z \in O_i}^{\infty} L_z^+.$$

■

Remark 2.

This last corollary simply says that maximal invariant control sets with nonempty interior in $\mathbb{R}_0^d$ under the action of a subsemigroup from Gl $(d, \mathbb{R})$ must be "cone shaped" in the sense that, when they contain a point, they necessarily contain the entire half line to which this point belongs. As seen above, this is basically a consequence of the linearity of this setup. Hence, controllability properties of linear semigroups S on $\mathbb{R}_0^d$ can be studied through their canonical action on the projective space $\mathbb{P}^{d-1}$. As we will see in subsequent sections, their growth behavior in $\mathbb{R}_0^d$ can also be studied via their projection on $\mathbb{P}^{d-1}$.

Remark 3.

The linear groups $G \subset$ Gl $(d, \mathbb{R})$ which act transitively on $\mathbb{R}_0^d$ were completely classified by Boothby and Wilson (1979). Transitive action on $\mathbb{R}_0^d$ implies, of course, transitive action on $\mathbb{P}^{d-1}$ or $\mathbf{S}^{d-1}$.

3.3. A Crucial Condition: The Orbits have Nonvoid Interior

After the broad discussion of Subsections 3.1 and 3.2, it should be clear that control sets properties are related to orbits properties. In fact, among other things,

we have shown that, given a semigroup $\mathcal{S}$ acting on a manifold M, int $Sx \neq \phi$ for all $x \in M$ implies that any maximal invariant control set C is necessarily closed and has a nonempty interior. This relationship is crucial to this thesis and will be expanded upon in this section. Nevertheless, for a complete picture, we first need to introduce the notions of (local) accessibility, (local) controllability, and transitivity. Since local accessibility and local controllability are properties of control systems that are not solely dependent on the semigroup $\mathcal{S}$ generated by the system but also on the path a point will follow under the action of $\mathcal{S}$, we start with the following definition:

Definition 1.

Assume one has a state space consisting of a manifold M and a subset $D \subset C(M)$, where C (M) represents the sets of all continuous maps from M to M. Then a trajectory or path starting at $x \in M$ and associated with $\mathcal{S}(D)$, the semigroup generated by D, is a sequence $g_n \circ \dots \circ g_1(x)$, with $g_i \in D$, $i \in \{1, \dots, n\}$, $n \in \mathbb{N}$.

Remark 1.

1) As already mentioned at the beginning of Subsection 3.1, discrete or continuous dynamical control systems are specific examples where the notion of path or trajectory is relevant. In these cases, the set $D \subset C(M)$ arises from the dynamics of the system.

   Also note that different systems on M may have the same (semi)group but different trajectories. Indeed, we may have two dynamics D_1 and $D_2 \subset C(M)$ with $D_1 \neq D_2$, but $\mathcal{S}(D_1) = \mathcal{S}(D_2)$.

2) When one deals with semigroups of continuous maps, i.e., $D \subset C(M)$, the group generated by $\mathcal{S}$, $\mathcal{G}(\mathcal{S})$, may not be defined. Indeed, some functions in $\mathcal{S}$ may not

be invertible. In the following (starting with Definition 2), whenever we will use the notion of a group arising from a dynamical system (and its associated semigoup), it will be understood that this group is assumed to be well defined. This is obviously the case if, in fact, $D \subset \mathrm{Diff}(M)$.

Definition 2.

A control system Σ is said to be controllable from $x \in M$ if $Sx = M$. It is locally controllable at $x \in M$ if, for any neighborhood U_x of x, the set $S_{U_x}x \cap U_x$ contains a neighborhood of x, where $S_{U_x}x$ denotes the trajectories of the semigroup S from x lying entirely within U_x.

The system Σ is said to be accessible from $x \in M$ if $Gx = M$ and locally accessible at $x \in M$ if, for any neighborhood U_x of x, $G_{U_x}x$ has nonvoid interior in U_x (in the topology of M), where $G_{U_x}x$ denotes the trajectories of the group $\mathcal{G}(S)$ from x staying entirely within U_x.

If any of the above concepts holds for all $x \in M$, we say that the system is controllable, locally controllable, accessible, etc.

Definition 3.

Let $\mathcal{G}$ be a group acting on a manifold M. Then $\mathcal{G}$ is said to act transitively at $x \in M$ if $Gx = M$. If this holds for all $x \in M$, we say that $\mathcal{G}$ acts transitively on M or simply that $\mathcal{G}$ is transitive on M.

Remark 2.

In fact, $\mathcal{G}$ acting transitively at $x \in M$ always implies that $\mathcal{G}$ is transitive on M. Indeed, for any two points $y, z \in M$, transitivity at x obviously implies that

$y, z \in Gx$ and conversely that $x \in Gy$ and $x \in Gz$. Hence, $y \in Gz$ and $z \in Gy$, which shows that $\mathcal{G}$ is transitive on M. From this it is also clear that accessibility from some $x \in M$ implies accessibility and that accessibility and transitivity are equivalent notions, the first referring to the paths and the second to the group.

Proposition 1.

The following relationships hold:

a) local accessibility implies accessibility and

b) (local) controllability implies (local) accessibility.

The converses of the above implications are false.

Proof.

a) Let $x \in M$ be arbitrary and U be a neighborhood of x. First we show that the definition of local accessibility implies that $G_U x \cap U$ contains a neighborhood of x. Indeed, there exists $g \in \mathcal{G}_U$ (defined to be the subset of $\mathcal{G}$ giving, from x, trajectories staying entirely in U) with a neighborhood of g (x), V_g, included in int $G_U x$. This implies that $x \in O \equiv \left[g^{-1}(V_g) \cap U\right]$. Hence, O is a neighborhood of x and $G_U x \cap U$ contains a neighborhood of x. Since, in this construction, the paths from x to any $y \in O$ never left U, $O \in G_U x$ and we are done.

Now we start the main argument. Since M is connected and Riemannian, for any $y \in M$, there is a path P_{xy} between x and y. Let U still be some neighborhood of x. Then $G_U x \cap U$ contains an open set containing x and so does $Gx \cap U$.
Therefore, there is $z_1 \in P_{xy} \cap Gx \cap U$ with $\rho(z_1, x) \equiv \epsilon_1 > 0$.
Now $z_1 \in Gx$ and since we have local accessibility at all $z \in M$, we can repeat the above argument to generate $z_2 \in Gz_1 \cap P_{xy}$ such that $\rho(z_1, z_2) \equiv \epsilon_2 > 0$. But P_{xy}

has finite length (Boothby (1986, pp. 187 ff.)). Hence, it suffices to show that $\sum_{n=1}^{\infty} \epsilon_n \geq L(P_{xy}) < \infty$ (where the length $L(P_{xy})$ of a path P_{xy} is defined as in Subsection 2.3) to prove that the sequence $\{z_k ; k \geq 1\}$ can be continued until the entire length of P_{xy} is covered, i.e., that, for some n, we have $y \in Gz_n$. Indeed, this means that $y \in Gz_n = Gz_{n-1} = \ldots = Gz_1 = Gx$.

Claim: The sequence $\{\epsilon_n ; n \geq 1\}$ can be chosen such that

$$\sum_{n=1}^{\infty} \epsilon_n \geq L(P_{xy}), \text{ x, y, and } P_{xy} \text{ arbitrary.}$$

Suppose that $\sum_{n=1}^{\infty} \epsilon_n = K < L(P_{xy})$ for some $\{\epsilon_n\}$ sequence. Then the sequence $\{z_k ; k \geq 1\}$ converges to some $z \in P_{xy}$ with $L(P_{xz}) < L(P_{xy})$. By local accessibility, Gz contains a neighborhood V_z of z. Since $z_k \longrightarrow z$ as $k \longrightarrow \infty$, for k large enough we must have $z_k \in V_z$, i.e., $z_k \in Gz$. But this also implies that $z \in Gz_k$ and hence $V_z \subset Gz_k$. From this it follows that the path from z to y, P_{zy}, intersects Gz_k and hence that the $\{z_k\}$ construction could be pursued beyond z. This proves our claim and completes the whole argument.

b) Trivial by definition.

To complete the proof of this proposition, it suffices to show that the converse of these two statements does not hold. This is done via the counterexamples below. ■

Example 1.

a) Let $M = \mathbb{R}^2_0$ and define $\{B_{r\varphi} ; r \in (0,\infty), \varphi \in (0, \frac{\pi}{2})\}$ to be a collection of matrices of the form

$$B_{r\lambda} \equiv \frac{1}{r}\begin{bmatrix} \cos(\theta+\varphi) & \sin(\theta+\varphi) \\ -\sin(\theta+\varphi) & \cos(\theta+\varphi) \end{bmatrix},$$

where θ is a fixed number in the interval $[\frac{\pi}{2}, \pi]$.

Let $\mathcal{S}$ be the semigroup generated by the collection $\{B_{r\varphi}\,;\, r \in (0,\infty), \varphi \in (0,\frac{\pi}{2})\}$. Then, for all $x \in \mathbb{R}^2_0$, $Sx = Gx = \mathbb{R}^2_0$, i.e., the system associated with $\mathcal{S}$ is accessible and controllable.

Nevertheless, this system in neither locally accessible nor locally controllable. Indeed, for all $x \in \mathbb{R}^2_0$, one can find a neighborhood of U of x such that $U \cap \mathcal{L}_x = \phi$ where, using polar coordinates,

$$\mathcal{L}_x \equiv \bigcup \{(r,\theta) \in \mathbb{R}^+ \times [0,2\pi) : \theta \notin [\theta_x - \tfrac{\pi}{2}, \theta_x + \tfrac{\pi}{2}]\} \text{ with } x = (r_x, \theta_x).$$

But the action of the matrices $B_{r\varphi}$ involves a rotation by an angle $\theta + \varphi$ with $\frac{\pi}{2} < \theta + \varphi < 3\frac{\pi}{2}$. This means that, from any $x \in U$, other points in U cannot be reached without using a path that intersects $\mathcal{L}$ and hence a path that leaves U.

In the above example, replacing $r \in (0,\infty)$ by $r \in (1,\infty)$ yields a system which is still accessible but not controllable.

Finally, if $\mathcal{S}$ is simply generated by the collection $\{B_{r\varphi}\,;\, r \in (1,\infty),\ \varphi \in (0,2\pi)\}$ with $B_{r\varphi} = \frac{1}{r}\begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix}$, the system is locally accessible but neither locally controllable nor controllable.

b) Because of our interest in $\mathbb{S}^{d-1}$ or $\mathbb{P}^{d-1}$, we give another series of examples on $\mathbb{P}^{d-1}$ (which also apply to $\mathbb{S}^{d-1}$) with $d = 2$.

With $M = \mathbb{P}^1$, let $\{A_\varphi\,;\, \varphi \in (0,\frac{\pi}{4})\}$ be a collection of rotation matrices with angle φ. Then it is clear that the system associated with $\mathcal{S}(\{A_\varphi\})$ is locally accessible

but not locally controllable. The same system also shows that we may have controllability without local controllability.

To show that accessibility does not imply local accessibility, it suffices to consider the semigroup $\mathcal{S}$ generated by the collection of rotation matrices $\{A_\varphi ; \varphi \in (\frac{\pi}{4}, \frac{\pi}{2})\}$ on $\mathbb{P}^1$.

Finally, we need to give an example for a system which is accessible (and even locally accessible) but not controllable. Since a rigorous description of example would be somewhat lengthy but involves only elementary linear algebra, we will describe the situation in an unformal way. First we will discuss the underlying argument on which this example is based.

Let $\{e_1, e_2\}$ be an orthogonal basis of $\mathbb{R}^2_0$ and consider the semigroup $\mathcal{S}$ (and its (canonical) action on $\mathbb{P}^1$, i.e., $g\, s \equiv g\, s\, |g\, s|^{-1}$, $s \in \mathbb{P}^1$ and $g \in \mathcal{S}$) generated by the following two matrices $A, B \in Gl(2, \mathbb{R})$:

A is a matrix with fixed eigenvalues $\lambda_1 > \lambda_2 > 0$ and corresponding eigenspaces $E_{\lambda_1} = \text{span}(e_1)$ and $E_{\lambda_2} = \text{span}(e_2)$ and

B is a matrix with fixed eigenvalues $\lambda_1' > \lambda_2' > 0$ and corresponding eigenspaces $E'_{\lambda_1'} = \text{span}(e_1 - e_2)$ and $E'_{\lambda_2'} = \text{span}(e_1 + e_2)$.

Note that $\mathcal{S} = \mathcal{S}(\{A, B\})$ (and $\mathcal{G} = \mathcal{G}(\{A, B\})$) contains matrices that are neither A nor B (because repeated matrix products between A and B will, among other things, generate matrices whose eigenvalues are not in $\{\lambda_1, \lambda_2, \lambda_1', \lambda_2'\}$). But we can still obtain sufficient insight into the action of $\mathcal{S}$ by simply examining the action of the matrices A and B.

The path of an element $s \in \mathbb{P}^1$ under the action of the matrix A will not cross the

eigenspaces $E_{\lambda_1} = \text{span}\,(e_1)$ and $E_{\lambda_2} = \text{span}\,(e_2)$ and, under the action of the matrix B, the path of an element $s \in \mathbb{P}^1$ will not cross the eigenspaces $E_{\lambda_1'} = \text{span}\,(e_1 - e_2)$ and $E_{\lambda_2'} = \text{span}\,(e_1 + e_2)$, the eigenspaces themselves remaining invariant. Moreover, the magnitude of the eigenvalues will determine the radial direction in which points in $\mathbb{P}^1$ will move: from E_{λ_2} to E_{λ_1} under the action of A and from $E_{\lambda_2'}$ to $E_{\lambda_1'}$ under the action of B. The picture on the next page (Figure 1) summarizes these facts. The single arrows ($\longrightarrow$) indicate the radial direction imposed by the matrix A (outside of E_{λ_1} and E_{λ_2}) while the double arrows ($\longrightarrow\!\!\!\!\!\rightarrow$) indicate the direction imposed by the matrix B (outside of $E_{\lambda_1'}$ and $E_{\lambda_2'}$).

Representing points in $\mathbb{P}^1$ by their angle in radians, this picture shows that, for example, points in $[3\frac{\pi}{4}, \pi]$ are trapped (under $\mathcal{S}$) in this same interval. Hence the system is not controllable. But, under $\mathcal{G}(\mathcal{S})$, the group generated by $\mathcal{S}$, all the arrows can be reversed at will and the system is (locally) accessible.

The above argument can then be used to construct a dynamical system with the desired features. Consider the following controlled difference equation on $\mathbb{P}^1$ (or S^1):

$$s_{n+1} = A_{\alpha\beta\theta}\, s_n \,/\, |A_{\alpha\beta\theta}\, s_n|,$$

where the matrix $A_{\alpha\beta\theta} = \begin{bmatrix} \alpha(\cos\theta)^2 + \beta(\sin\theta)^2 & (\alpha-\beta)\sin\theta\cos\theta \\ (\alpha-\beta)\sin\theta\cos\theta & \beta(\cos\theta)^2 + \alpha(\sin\theta)^2 \end{bmatrix}$,

with $\alpha > \beta > 0$ and $\theta \in (3\frac{\pi}{4}, \pi)$.

One can then verify that all the $A_{\alpha\beta\theta}$ matrices have eigenvalues $\lambda_1 = \alpha > \beta = \lambda_2$ with corresponding eigenspaces

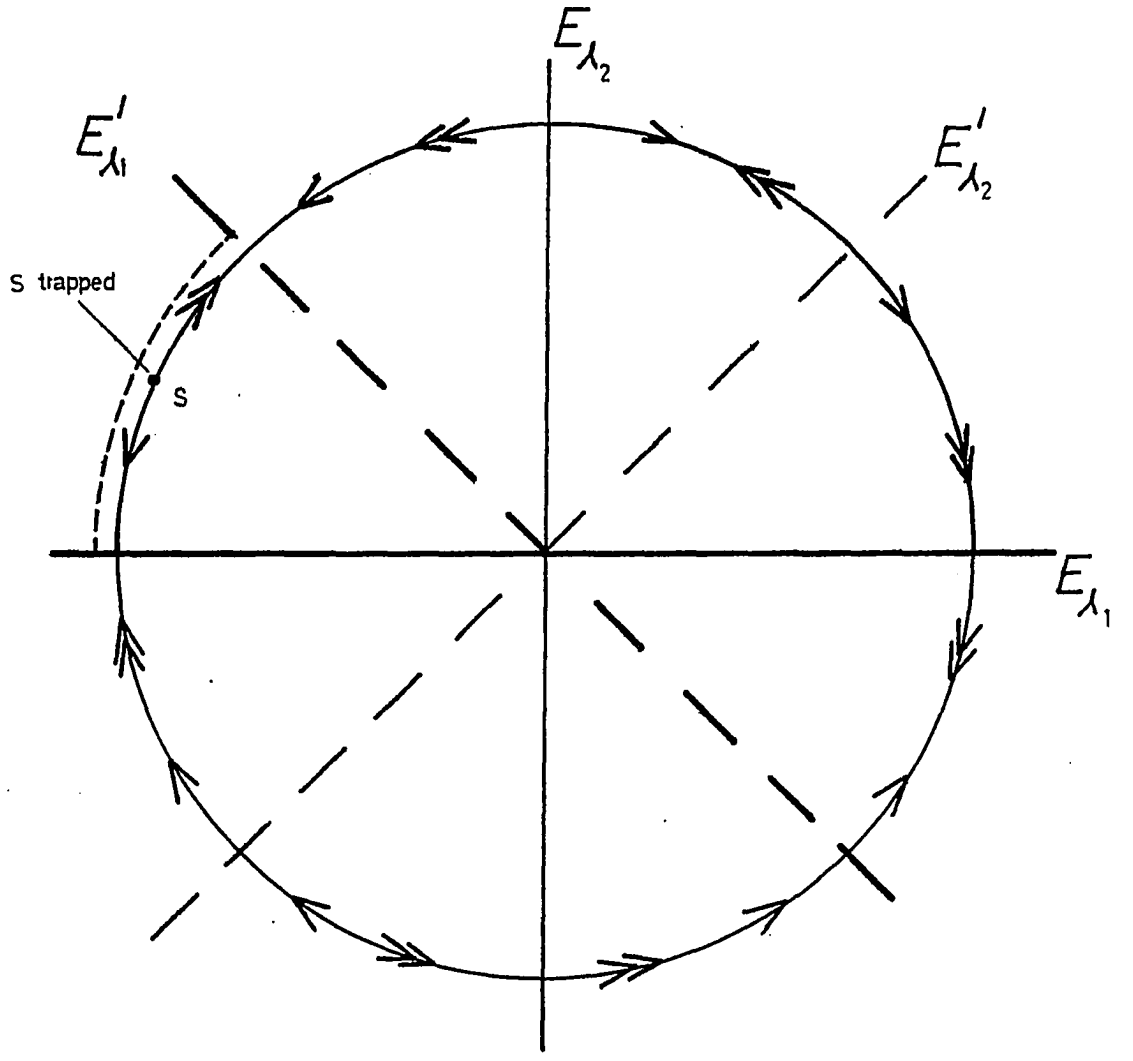


Figure 3.1. Example of a (locally) accessible but not controllable system

$$E_{\lambda_1} = \text{span}([\cos\theta, \sin\theta]') \text{ and}$$

$$E_{\lambda_2} = ([\sin\theta, -\cos\theta]'),$$

where $[x, y]' = \begin{bmatrix} x \\ y \end{bmatrix}$.

This setup is similar to the one discussed above. In fact the collection of matrices $A_{\alpha\beta\theta}$ is used to ensure that, by varying α, β, and θ at will, all points $s \in (3\frac{\pi}{4}, \pi)$ satisfy $Ss \subset (3\frac{\pi}{4}, \pi)$. The matrices A and B above simply correspond to the extreme choices $A_{\lambda_1\lambda_2\pi}$ (= A) and $A_{\lambda_1'\lambda_2'(\frac{3}{4}\pi)}$ (= B). It is then easy to see that, similarly to the situation described above for the semigroup $\mathcal{S}(\{A, B\})$, we have that $s \in (3\frac{\pi}{4}, \pi)$ implies $Ss \subset (3\frac{\pi}{4}, \pi)$, i.e., the system is not controllable while, $Gs = \mathbb{P}^1$ (or $\mathbb{S}^1$) and we have (local) accessibility.

Even though the notion of (local) controllability will not play an important role in this thesis, we will, for sake of completeness, say a few more words about it, before stating the main result of this section and its consequences.

The question still to be addressed concerns the relationship between the notions of controllability and local controllability. Both of these deal with the precise reachability of points and, therefore, their relationship will involve properties of the negative orbit $S^{-}x$, i.e., of the points that can be controlled to $x \in M$. In the continuous case, the integrability of the Lie algebra of the vector fields giving the system's dynamics is sufficient to guarantee that, for any neighborhood U_x of x, the sets $S_{U_x}x \cap U_x$ and $S^{-}_{U_x}x \cap U_x$ have nonvoid interior (in the maximal integral manifold through x). In control sets C with int $C \neq \phi$, this implies controllability in

int C, i.e., int C $\subset$ Sx for all x $\in$ int C (and hence int C = Sx by the C–invariance of int C in the continuous case). Similarly, local controllability in M can be shown to imply controllability in M. In the discrete time case, a similar situation arises in the sense that we will need both of the conditions int Sx $\neq \phi$ and int $S^{-}x \neq \phi$ (one does not imply the other, see Example 3.1.6 (a)) to relate the notions of local controllability and controllability. We will formulate and prove the next result in three steps.

Proposition 2.

Local controllability implies $\overline{Sx} = M$ for all x $\in$ M.

Proof.

Assume that $\overline{Sx} \subset M$ but $\overline{Sx} \neq M$ for some x $\in$ M and pick y $\in \partial \overline{Sx}$. Local controllability at y means that, for any neighborhood U_y of y, the set $U_y \cap S_{U_y} y$ contains a neighborhood of y, say V_y. Since $y \in \partial \overline{Sx} \neq M$, there must exist $z \in (M \setminus \overline{Sx}) \cap V_y$ and $g \in \mathcal{S}$ such that g (y) = z. Let W_z be an open neighborhood of z such that $W_z \subset V_y$ and $W_z \cap \overline{Sx} = \phi$. Then $g^{-1}(W_z)$ is an open neighborhood of y and, therefore, $g^{-1}(W_z)$ contains points of Sx. Hence, W_z also contains points of Sx. This contradicts the choice of W_z and so, $\overline{Sx} = M$.

■

Lemma 1.

If int $S^{-}x \neq \phi$ for all x $\in$ M, then $\overline{Sx} = M$ if and only if Sx = M.

Proof.

Pick y $\in$ M. Then int $S^{-}y \neq \phi$ and $\overline{Sx} = M$ imply that there exists z $\in$ Sx $\cap$ int $S^{-}y$, i.e., y $\in$ Sx.

■

Corollary 1.

If int $\bar{Sx} \neq \phi$ for all $x \in M$, then local controllability implies controllability.

Proof.

Immediate from Proposition 2 and Lemma 1.

■

Remark 3.

Under the stronger assumption of negative local controllability, i.e., the assumption that, for all $x \in M$ and all open neighborhood U_x of x, the set $\overline{S_{U_x}x} \cap U_x$ contains a neighborhood of x, we can mimic the proof of Proposition 1 (a) for the semigroup $\mathcal{S}$. In this case, local controllability implies that one can follow each path P_{xy} from x to y without deviating from the path by more than a prescribed $\epsilon > 0$. This property is called global asymptotic tracking.

In Remark 2, we have stated that accessibility at one point always implies accessibility. The last part of Example 1 (b) (on $\mathbb{P}^1$, not on $\mathbb{S}^1$) shows that the same is not true for controllability. Indeed, going back to the pictorial representation of the behavior of the system (Figure 1), one can see that the system is controllable from the points in $[\frac{\pi}{4}, \frac{\pi}{2}]$ while, as previously discussed, this is not true from the points in $[3\frac{\pi}{4}, \pi]$.

After this review of the notions of (local) accessibility, transitivity, and (local) controllability, we are now ready to complete the overall picture by discussing the deterministic assumption which is central to this thesis:

$$\text{int } Sx \neq \phi \text{ for all } x \in M.$$

This is an assumption about the "richness" of the orbits under the semigroup $\mathcal{S}$.

First note that, as shown by Example 3.1.6 (a), this condition implies neither (local) controllability nor local accessibility. This explains why the notion of (local) controllability is only marginal in this thesis and why control sets (defined only up to closure of the orbits) will play an important role in the following sections . The following lemma shows the relationship between this assumption and transitivity (and hence, by Remark 2, accessibility).

Lemma 2.

Let $\mathcal{S}$ be a semigroup acting on M. If, for all $x \in M$, int $Sx \neq \phi$, then $\mathcal{G}(\mathcal{S})$ acts transitively on M, i.e., $Gx = M$ for all $x \in M$.

On the other hand, transitivity does not imply int $Sx \neq \phi$ for all $x \in M$.

Proof.

Let $x \in M$ be arbitrary, U_x be any neighborhood of x, and pick $y \in$ int $Sx \neq \phi$. Then $y = g(x)$ for some $g \in \mathcal{S}$. Hence $x \in g^{-1}$ (int Sx) $\subset Gx$ and therefore $Gx \cap U_x \neq \phi$. An examination of the proof of Proposition 1 (a) shows that this is enough to conclude that we have accessibility and therefore that $\mathcal{G}$ is transitive on M (see Remark 2). That transitivity does not imply int $Sx \neq \phi$ for all $x \in M$ can be seen from an example found in Jacubczyk and Sontag (1988, Remark 5.1). ∎

When $\mathcal{S} \subset Gl(d, \mathbb{R})$ and $M = \mathbb{P}^{d-1}$, the results collected above as well as in Subsection 3.1 give us the following key result, which is a slight enhancement of Proposition 4.1 in San Martin and Arnold (1986).

Theorem 1.

Let $\mathcal{S} \subset Gl(d, \mathbb{R})$ be a semigroup whose canonical action on $\mathbb{P}^{d-1} \subset \mathbb{R}^d$ satisfies the condition

$$\text{int } Ss \neq \phi \quad \text{for all } s \in \mathbb{P}^{d-1}.$$

Then there exists a unique maximal invariant control set C for $\mathcal{S}$ on $\mathbb{P}^{d-1}$. C is closed (and hence Borel), with nonempty interior, and given by

$$C = \bigcap \{\overline{Ss} ; s \in \mathbb{P}^{d-1}\} \neq \phi.$$

Proof.

Since int $Ss \neq \phi$ for all $s \in \mathbb{P}^{d-1}$, $\mathcal{G}(\mathcal{S})$ is a Lie subgroup of $Gl(d, \mathbb{R})$ which acts transitively on $\mathbb{P}^{d-1}$. The result then follows from Proposition 4.1 in San Martin and Arnold (1986), upon noting that the condition int $\mathcal{S} \neq \phi$ in $\mathcal{G}$ ($\mathcal{G}$ with its Lie group topology and acting transitively on $\mathbb{P}^{d-1}$) stated in this result is only used to conclude that int $Ss \neq \phi$ for all $s \in \mathbb{P}^{d-1}$. That int $C \neq \phi$ follows trivially from $Sx \subset \overline{C} = C$ for all $x \in C$.

■

Remark 4.

Assessing the uniqueness of the (deterministic) maximal invariant control set associated with a stochastic difference or differential equation (see Subsection 4.1) is an important step in the study of the stability properties of such an equation. In the continuous time case, results similar to Theorem 1 can be found in Arnold et al. (1986a). Here we used the discrete time version of one of these results, described in San Martin and Arnold (1986).

As hinted to in the proof of Theorem 1, one way to ensure that int $Sx \neq \phi$ for all $x \in \mathbb{R}_0^d$ (or $\mathbb{S}^{d-1}$, or $\mathbb{P}^{d-1}$), $\mathcal{S} \subset Gl(d, \mathbb{R})$, is to require that int $\mathcal{S} \neq \phi$ in some Lie subgroup of $Gl(d, \mathbb{R})$ containing $\mathcal{G}(\mathcal{S})$ and acting transitively on $\mathbb{R}_0^d$ (or $\mathbb{S}^{d-1}$, or $\mathbb{P}^{d-1}$). This follows from a simple argument based on Theorem 2.2.1. Nevertheless, one can use another condition due to Jakubczyk and Sontag (1988), which has the beauty of relating the assumption int $Sx \neq \phi$ to a Lie Algebra condition, as in the continuous time case. Before stating this result, we need to set up some notation. For a more complete discussion, see Jakubczyk and Sontag (1988).

Consider the controlled difference equation

$$x_{n+1} = f(x_n, u),$$

where U is a subset of $\mathbb{R}^m$ satisfying $0 \in U \subset \overline{\text{int } U}$ and $f : M \times U \longrightarrow M$ satisfies the condition that $f_u \equiv f(., u) : M \longrightarrow M$ is a C^∞ diffeomorphism for each u.

Write $f_{u_k \ldots u_1} \equiv f_{u_k} \circ \ldots \circ f_{u_1}$ and, in order to use the inverse of f_u, f_u^{-1}, also write $f^{\epsilon_k \ldots \epsilon_1}_{u_k \ldots u_1} \equiv f^{\epsilon_k}_{u_k} \circ \ldots \circ f^{\epsilon_1}_{u_1}$, where each of $\epsilon_1, \ldots, \epsilon_k$ takes a value of ± 1. For $x \in M$ and $u \in U$, define the following two families of vector fields

$$X_u^+(x) = \frac{\partial}{\partial v}\Big|_{v=0} f_u^{-1} \circ f_{u+v}(x) \text{ and}$$

$$X_u^-(x) = \frac{\partial}{\partial v}\Big|_{v=0} f_u \circ f_{u+v}^{-1}(x).$$

Finally, given a vector field Y and a control value u, define the operator Ad_u by

$$\left[\mathrm{Ad}_u\, Y\right](x) = (d\, f_u(x))^{-1}\, Y(f_u(x)),$$

where $d\, f_u$ denotes the differential of f_u with respect to x. Using the diffeomorphism Ad_u, we may also define

$$\left[\mathrm{Ad}^{\epsilon_k \cdots \epsilon_1}_{u_k \ldots u_1} Y\right](x) = \left[d\, f^{\epsilon_k \cdots \epsilon_1}_{u_k \ldots u_1}(x)\right]^{-1} Y\left[f^{\epsilon_k \cdots \epsilon_1}_{u_k \ldots u_1}(x)\right].$$

Using the abbreviation $\mathrm{Ad}^k_0\, Y$ for $\mathrm{Ad}_{0\ldots0}\, Y$ with $u = 0$ repeated k times, if $k > 0$, for $\mathrm{Ad}^{-1\ldots-1}_{0\ \ldots0}\, Y$, if $k < 0$, and $\mathrm{Ad}^0_0\, Y = Y$, we get that

$$\left[\mathrm{Ad}^k_0\, X^+_u\right](x) = \frac{\partial}{\partial v}\Big|_{v=0} f_0^{-k} \circ f_0^{-1} \circ f_{u+v} \circ f_0^k(x) \text{ and, more generally, that}$$

$$\left[\mathrm{Ad}_{u_k \ldots u_1} X^+_{u_0}\right](x) = \frac{\partial}{\partial v}\Big|_{v=0} f^{-1}_{u_k \ldots u_1} \circ f^{-1}_{u_0} \circ f_{u_0+v} \circ f_{u_k \ldots u_1}(x).$$

<u>Theorem 2.</u>

Define the families of vector fields Γ^+, Γ^-, and Γ by

$$\Gamma^+ \equiv \left\{\mathrm{Ad}_{u_k \ldots u_1} X^+_{u_0};\ k \geq 0,\ u_0, \ldots, u_k \in U\right\},$$

$$\Gamma^- \equiv \left\{\mathrm{Ad}^{-1\ldots-1}_{u_k \ldots u_1} X^-_{u_0};\ k \geq 0,\ u_0, \ldots, u_k \in U\right\},$$

$$\Gamma \equiv \left\{\mathrm{Ad}^{\epsilon_k \cdots \epsilon_1}_{u_k \ldots u_1} X^{\sigma}_{u_0};\ k \geq 0,\ u_0, \ldots, u_k \in U,\ \epsilon_1, \ldots, \epsilon_k = \pm 1,\ \sigma = \pm\right\}.$$

Let $\dim \Gamma(x)$ (respectively $\dim \Gamma^+(x)$, $\dim \Gamma^-(x)$) represent the dimension of the linear space spanned by the vectors in Γ (respectively Γ^+, Γ^-) evaluated at x. Let Lie (Δ) be the Lie algebra generated by some collection of vector fields Δ, and let

Lie (Δ) (x) denote the linear space of tangent vectors at x given by the vector fields in Lie (Δ).

Then, for any C^{∞} system, we have that

a) int $Sx \neq \phi$ for all $x \in M^d$ if and only if, for all $x \in M^d$, dim $\Gamma^+(x) = d$ (which is equivalent to dim Lie $(\Gamma^+)(x) = d$),

b) int $S^-x \neq \phi$ for all $x \in M^d$ if and only if, for all $x \in M^d$, dim $\Gamma^-(x) = d$ (which is equivalent to dim Lie $(\Gamma^-)(x) = d$) and

c) the system's group acts transitively on M^d if and only if, for all $x \in M^d$, either dim $\Gamma(x) = d$ (which is equivalent to dim Lie $(\Gamma)(x) = d$).

Proof.

See Jakubczyk and Sontag (1988, Theorem 4.2).

■

Remark 5.

Note that Theorem 2 gives symmetrical conditions for both int $Sx \neq \phi$ and int $S^-x \neq \phi$. The latter will not be used in the subsequent chapters but it is worth mentioning that all the results involving int Sx previously given have their obvious counterparts for int S^-x. Moreover, recall that, when, for all $x \in M$, both int $Sx \neq \phi$ and int $S^-x \neq \phi$ are satisfied, local controllability implies controllability. The discrete time situation then closely parallels the continous time case.

Remark 6.

The conditions stated here to ensure that int $Sx \neq \phi$ or int $S^-x \neq \phi$ have the additional practical advantage that they are relatively easy to check. Recall nevertheless that many of our results only require weaker assumptions (like int $\overline{Sx} \neq \phi$) which

automatically follow from int $Sx \neq \phi$ (see Subsection 3.1).

In the remainder of this thesis, we will need the existence of a unique, closed maximal invariant control set with nonempty interior. As shown above and similarly to the continuous time case situation, the condition int $Sx \neq \phi$ for all $x \in M$ implies transitivity and will (at least for linear systems projected on $\mathbb{P}^{d-1}$) give us the existence of such a unique, closed maximal invariant control set with nonempty interior (via Theorem 1). Some of the results were shown to require the additional assumption that int $S^{-}x \neq \phi$ for all $x \in M$ (e.g., local controllability implies controllability or int $Sx = \text{int } C$ for all $x \in \text{int } C$) but these results are not needed hereafter. Again recall that, in the continuous case, both int $Sx \neq \phi$ and int $S^{-}x \neq \phi$ (for all $x \in M$) follow from the same Lie algebra condition while, in the discrete time case, Theorem 2 clearly exhibits separate assumptions for the positive and negative orbits. Recall Example 3.1.6 (a), which showed that, in the discrete time case, one can have int $Sx \neq \phi$ for all $x \in M$ (i.e., dim Lie $\Gamma^{+}(x) = d$ for all $x \in M$) but int $S^{-}x = \phi$ for some $x \in M$ (i.e., dim Lie $\Gamma^{-}(x) < d$ for some $x \in M$). In the continuous time case and without the Lie algebra condition, such a behavior is also possible (see the example in Arnold and Kliemann (1987, Remark 3.3)).

4. STOCHASTIC DIFFERENCE EQUATIONS

4.1. Basic Setup

Section 4 is devoted to the existence and uniqueness of an invariant probability measure for the pair process $\{(x_n, \xi_n)\}$. We are considering the general stochastic difference equation

$$x_{n+1} = f(x_n, \xi_n) \qquad n \geq 0,$$

where

x_n takes values in a C^∞ manifold M,

$(\Omega, \mathcal{F}, P)$ is some underlying probability space,

$\{\xi_n ; n \in \mathbb{N}\} = \{\xi_n(\omega) ; n \in \mathbb{N}\}$ is a Feller time homogenous Markov chain taking values in a C^∞ manifold W, and

$f : M \times W \to M$ is jointly measurable.

Further assumptions on the setup are to be made. If $\{\xi_n ; n \in \mathbb{N}\}$ is simply a sequence of iid random variables, the sequence $\{x_n ; n \in \mathbb{N}\}$ is itself an M valued Markov chain. For the more general case, we need the following result.

Proposition 1.

Let $(\Omega, \mathcal{F}, P)$ be a probability space and consider, on M, the stochastic difference equation

$$x_{n+1} = f(x_n, \xi_n(\omega)), \quad n \geq 0, \; x_0 = x_0(\omega) \text{ and } f \text{ jointly measurable},$$

where $\{\xi_n ; n \in \mathbb{N}\} = \{\xi_n(\omega) ; n \in \mathbb{N}\}$ is a Markov chain with state space $(W, \mathcal{B}(W))$. Assume that the random initial value x_0 is independent of the process $\{\xi_n ; n \in \mathbb{N}\}$. Then, given some fixed initial value $x_0(\omega) = y$, the solution at the n^{th} step $x(n, y, \omega)$ is a $\sigma(\xi_k ; 0 \leq k \leq n-1) \times \mathcal{B}(M)$ measurable function of from $\Omega \times M \longrightarrow M$. Also, for a random initial value $x_0(\omega)$, $x(n, x_0(\omega), \omega)$ is a $\sigma(x_0, \xi_k ; 0 \leq k \leq n-1)$ measurable function of from $\Omega \longrightarrow M$.

Moreover, for a random initial value $x_0(\omega)$, $\{z_n\} \equiv \{(x_n, \xi_n)\}$ is a time homogeneous Markov chain with transition probabilities given, for $B \in \mathcal{B}(W \times M)$, by

$$\bar{\mu}(n, z, B) \equiv P((x(n, z_1, \omega), \xi_n(\omega)) \in B \mid \xi_0(\omega) = z_2),$$

where $z = (z_1, z_2)$, $z_1 \in M$ and $z_2 \in W$.

We also have that

$$\begin{aligned}
&\bar{\mu}(n, (x_0(\omega), \xi_0(\omega)), B) \\
&\quad \equiv P((x_n(\omega), \xi_n(\omega)) \in B \mid (x_0(\omega), \xi_0(\omega))) \\
&\quad = P((x(n, x_0(\omega), \omega), \xi_n(\omega)) \in B \mid \xi_0(\omega)).
\end{aligned}$$

Finally, if $\{\xi_n ; n \in \mathbb{N}\}$ is Feller, so is $\{(x_n, \xi_n)\}$.

Proof.

See Bunke (1972, Proposition 6.1, p. 138). The proof given there is concerned with the continuous time case but immediately applies to the discrete time case. A formulation of this result can also be found in Arnold and Kliemann (1983, Lemma 2.1), where the proof that the process $\{(x_n, \xi_n)\}$ is Feller can be found.

■

Remark 1.

Upon recalling the recursive definition $f_{\xi_k \dots \xi_0}(x) = f(f_{\xi_{k-1} \dots \xi_0}(x), \xi_k)$ introduced in Subsection 2.4, $P((x(n, z_1), \xi_n) \in B \mid \xi_0 = z_2)$ reads

$$P((f_{\xi_{n-1} \dots \xi_m}(z_1), \xi_n) \in B \mid \xi_0 = z_2)).$$

This formulation shows that the map f only influences the first component of the joint Markov chain and the structure of the dependency on the the ξ_n path and on the initial value $z = (z_1, z_2)$.

Now we are interested in the ergodic behavior of the pair process $\{(x_n, \xi_n)\}$, i.e., we wish to obtain the existence and uniqueness of an invariant probability measure for the pair process $\{(x_n, \xi_n)\}$. For systems in $\mathbb{R}^d$ with $\{\xi_n ; n \in \mathbb{N}\}$ being an iid sequence, this has been done by Meyn (1989) and Meyn and Caines (1988) but, for this more general setup involving a pair process, we need further assumptions on the Markov chain $\{\xi_n ; n \in \mathbb{N}\}$ and the dynamics f. Hence, our basic working assumptions are given here below.

On the general setup:

1) M^d and W^k are C^∞ manifolds and

2) the map $f : M \times W \longrightarrow M$ is jointly C^1.

On $\{\xi_n ; n \in \mathbb{N}\}$:

3) $x_0 = x_0(\omega)$ is independent of $\{\xi_n ; n \in \mathbb{N}\}$.

4) $\{\xi_n ; n \in \mathbb{N}\}$ is an ergodic (i.e., positive and weakly recurrent) W–valued Feller Markov chain with invariant probability measure $\pi_\xi \in \mathcal{P}(W)$, the probability measures on the space $(W, \mathcal{B}(W))$.

5) With supp $(\pi_\xi) \equiv Q \subset W$ and using $\mu(\xi, B)$ to denote the one-step transition probability from $\xi \in W$ to $B \in \mathcal{B}(W)$, $\mu(\xi, B) > 0$ for all $\xi \in$ int Q and all sets $B \in \mathcal{B}(W)$ such that $\pi_\xi(B) > 0$.
Moreover, $\pi_\xi(\text{int } Q) = \pi_\xi(Q)$.

6) In the decomposition of the one–step transition probability $\mu(\xi, .)$ into its absolutely continuous and singular components (with respect to the volume element m_W on W),

$$\mu(\xi, B) = \int_B g(\xi, y)\, m_W(dy) + \mu_s(\xi, B)$$

defined for $B \in \mathcal{B}(W)$ and $\xi \in W$, the density function $g(\xi, .)$ is, for all $\xi \in$ int Q, strictly positive on int Q, zero on $(\text{int } Q)^c$, and the map g is jointly lower semi–continuous.
Moreover, either $\mu_s(\xi, .) \equiv 0$ for all $\xi \in Q$ or $\{\xi_n\}$ is a stationary process.

On the dynamics:

7) For every pair $(x, \zeta) \in M \times$ int Q, there exist a time $n_{x\zeta}$ and an open set $O_{x\zeta} \subset M$ (the subscripts indicating a dependence on the points x and ζ) such that $m_M(O_{x\zeta} \cap .) \ll P(x_{n_{x\zeta}} \in . \mid x_0 = x, \xi_0 = \zeta)$, where we use $P(x_{n_{x\zeta}} \in . \mid x_0 = x, \xi_0 = \zeta) \equiv P((x_{n_{x\zeta}}, \xi_{n_{x\zeta}}) \in . \times W \mid x_0 = x, \xi_0 = \zeta)$,

and where m_M denotes the volume element on M.

8) The deterministic semigroup generated by the associated nonrandom dynamical system $x_{n+1} = f(x_n, u_n)$, $u_n \in \text{int } Q$, admits a unique invariant control set $C \subset M$ satisfying $C = \overline{C}$ (hence C is maximal and Borel) and int $C \neq \phi$.

Let us now discuss these assumptions.

Assumption 1.

Actually, M and W need not be C^∞ manifolds for most of the arguments in Section 4 to be true (C^1 is good enough), but C^∞ manifolds will be used in connection with diffeomorphisms when referring to Theorem 3.3.2. Also, some results will require the additional hypotheses that M and/or C and/or W be compact. When applicable, this will be specified.

Assumption 2.

The assumption that f be jointly C^1 is in fact too strong for many of the results hereafter. It is usually enough to have $f(., \xi)$ to be C^1 and $f(x, .)$ to be continuous. Also recall that Section 3 results only required $f(., \xi)$ to be continuous, except for Proposition 3.1.6 (which required the map $f(., \xi)$ to be C^1) and Proposition 3.1.10. Also, Theorem 1 below (used to obtain a condition ensuring the validity of Assumption 7 above) does require f to be jointly C^1. Nevertheless, Theorem 3.3.2 which can be used to verify both Assumptions 7 and 8 (as discussed later in this section) does require $f(., \xi)$ to be a diffeomorphism for all $\xi \in \text{int } Q$.

Assumption 3.

This assumption is simply used to guarantee that the pair process $\{(x_n, \xi_n)\}$ is Markov and Feller (see Proposition 1).

Assumption 4.

In order to obtain ergodicity for the pair process $\{(x_n, \xi_n)\}$, one must obviously be able to assert the ergodicity of the $\{\xi_n ; n \in \mathbb{N}\}$ Markov chain alone. As often found in the relevant literature, one could also require that the chain $\{\xi_n ; n \in \mathbb{N}\}$ be stationary. Nevertheless, this stationarity requirement is not necessary for the developments in Subsection 4.2 and will not be made here.

Assumption 5.

The assumption "$\pi_\xi(B) > 0$ implies $\mu(\xi, B) > 0$ for all $\xi \in$ int Q" replaces the support theorem, which is used in the continuous time case. It is a critical hypothesis to ensure a connection between the paths of the associated deterministic control system introduced in Assumption 8 and the transition probabilities of the stochastic pair process. Indeed, at every step, the associated deterministic control system allows the use of any control value in int Q and, without this assumption, this could result in a deterministic system which is much "richer" (in terms of paths) than its stochastic counterpart.

Note that, in fact, the first statement in Assumption 5 is implied by Assumption 6. This redundancy was deliberate in order to stress the importance of this assumption relating the possible paths of the associated deterministic control system to the stochastic paths and to enable easy references to this fact.

The additional assumption that $\pi_\xi(\text{int } Q) = \pi_\xi(Q)$ will be needed in Theorem 1 below. Moreover, under this last assumption, the control values for the associated deterministic control system (in Assumption 8) can be taken from the open set int Q. Even though this aspect is not crucial, we gain consistency with the requirement imposed in Subsection 3.1 that control values live in an open subset of $\mathbb{R}^k$.

Assumption 6.

The lower semi–continuous densities for the one step transition probabilities $\mu(\xi, .)$ will be used (via Theorem 1) to ensure that weak stochastic controllability (see Assumption 7) from one point also implies weak stochastic controllability from a whole neighborhood of that point. This will be needed because, for Markov processes on continuous state spaces, one cannot normally guarantee a positive probability of hitting single points or sparse sets.

Also note that, by definition of an invariant measure and since $\{\xi_n\}$ is Feller, $Q \equiv \text{supp}(\pi_\xi)$ is necessarily an invariant set for the process $\{\xi_n\}$. Moreover, the existence of a strictly positive (lower semi–continuous) density (with respect to m_W) on int Q for the one step transition probabilities $\mu(\xi, .)$ necessarily implies that $m_W(Q \cap .) \ll \pi_\xi$, i.e., that the volume element on W restricted to Q is absolutely continuous with respect to the unique invariant probability measure for the $\{\xi_n\}$ process. Indeed, if $m_W(B \cap Q) > 0$, $B \in \mathcal{B}(W)$, then

$$\pi_\xi(B) \geq \int_W \int_{B \cap Q} g(\xi, y)\, m_W(dy)\, \pi_\xi(d\xi) > 0.$$

The assumption that $\mu_s(\xi, .) \equiv 0$ (for all $\xi \in Q$) is used to ensure that, if $B \in \mathcal{B}(W)$ satisfies $\pi_\xi(B) = 0$, then $\mu(\xi, B) = 0$ for all $\xi \in Q$. Indeed, if $\pi_\xi(B) = 0$ (and hence

$m_W(B \cap Q) = 0)$, then we have $\mu(\xi, B) = \int_{B \cap Q} g(\xi, y)\, m_W(dy) = 0$, for all $\xi \in Q$. In particular, since we assume $\pi_\xi(\partial Q) = 0$, the set int Q is invariant for $\{\xi_n\}$. This fact will be implicitly used in several proofs given in Subsection 4.2. Also note that, under the hypothesis $\mu_s(\xi, .) \equiv 0$ for all $\xi \in Q$,

$$\pi_\xi(B) = \pi_\xi(B \cap Q) = \int_Q \int_{B \cap Q} g(\xi, y)\, m_W(dy)\, \pi_\xi(d\xi),$$

which shows that $\pi_\xi \ll m_W(Q \cap .)$. From this it follows that $\pi_\xi(\text{int } Q) = \pi_\xi(Q)$ and that $\pi_\xi(B) > 0$ implies $\int_B g(\xi, y)\, m_W(dy) = \mu(\xi, B) > 0$ for all $\xi \in Q$. Hence, under the assumption $\mu_s(\xi, .) \equiv 0$ for all $\xi \in Q$, Assumption 5 directly follows from Assumption 6.

Stationarity is another way to trivially ensure that $\pi_\xi(B) = 0$ implies $\mu(\xi, B) = 0$ for all $\xi \in Q$ (with $P(\xi_0 \in Q) = 1$) and that $\pi_\xi(B) > 0$ implies $\mu(\xi, B) > 0$ for all $\xi \in Q$ (this last statement being part of Assumption 5). In Subsection 4.2, we will use the condition $\mu_s(\xi, .) \equiv 0$ for all $\xi \in Q$ since this permits us to establish results involving the convergence in distribution to unique invariant probability measures for both the $\{\xi_n\}$ and $\{(x_n, \xi_n)\}$ processes. Clearly, all our results will remain true under the alternative hypothesis that $\{\xi_n\}$ is a stationary process. We then only have to deal with the convergence in distribution (to a unique invariant probability measure) of the pair process $\{(x_n, \xi_n)\}$. This would in fact simplify several of our arguments.

Assumption 7.

This assumption is called weak stochastic controllability (in the M component) by Meyn and Caines (1988) and will be further discussed later in this section.

Assumption 8.

This assumption summarizes all the characteristics required from the associated deterministic control system for a fruitful study of the behavior of the stochastic pair process $\{(x_n, \xi_n)\}$ via the tools of geometric control theory. Under our previous assumptions, the semigroup in question has the form

$$S = \left\{ f_{\xi_n \ldots \xi_0} \, ; \, \xi_i \in \operatorname{supp}(\pi_\xi) \, , i = 0, \ldots, n \, ; \, n \in \mathbb{N} \right\}.$$

The remainder of this section will be devoted to a more in depth discussion of Assumptions 7 and 8.

On a compact manifold, Proposition 3.1.12 guarantees the existence of a maximal invariant control sets in the closure of each orbit Sx. Moreover (and this fact does not require compactness), the same proposition tells us that, if $C \equiv \bigcap \{\overline{Sx} \, ; \, x \in M\} \neq \phi$, then the set C is automatically a maximal invariant control set. Furthermore, Corollary 3.1.2 tells us that, if int $C \neq \phi$, then C must be the unique maximal invariant control set and that, if a unique maximal invariant control set exists on a compact manifold, then this maximal invariant control set must be of the form $\bigcap \{\overline{Sx} \, ; \, x \in M\} \neq \phi$.

As we have discussed in detail at the end of Subsection 3.3 (see Theorem 3.3.1), when working with the projection on $\mathbb{P}^{d-1}$ of a linear system on $\mathbb{R}^d_0$, the assumption int $Ss \neq \phi$ for all $s \in \mathbb{P}^{d-1}$ guarantees that all the conditions of Assumption 8 are met. In particular, the unique maximal invariant control set is necessarily of the form $\bigcap \{\overline{Ss} \, ; \, s \in \mathbb{P}^{d-1}\} \neq \phi$.

Now the assumption int $Sx \neq \phi$ for all $x \in M$ is also related to Assumption 7. Nevertheless, in order to pursue along these lines, we need to introduce the notion of generalized controllability matrix and expand on some results obtained by Meyn and Caines (1988).

Consider the stochastic dynamical system

$$x_{n+1} = f(x_n, \xi_n), \quad n \geq 0,$$

where $f : M \times W \longrightarrow M$ is C^1 (i.e., continuously differentiable in both variables). For this system, the value of the stochastic orbit of a point $x \in M$ at time k will be denoted by $S^k_{x\,\xi_0} \equiv f_{\xi_{k-1}\cdots\xi_0}(x)$, $S^0_{x\,\xi_0} = x$. We can then give the following definition:

Definition 1.

For $x \in M$ and a nonrandom sequence $\{\xi_k ; k \geq 0\}$, let $\{A_k ; k \geq 0\}$ and $\{B_k ; k \geq 0\}$ denote the sequences of Jacobian matrices defined by

$$A_k = A_k(x, \xi_0, \ldots, \xi_k) \equiv \left[\frac{\partial f}{\partial x}\right]_{(S^k_{x\,\xi_0}, \xi_k)} \text{ and}$$

$$B_k = B_k(x, \xi_0, \ldots, \xi_k) \equiv \left[\frac{\partial f}{\partial \xi}\right]_{(S^k_{x\,\xi_0}, \xi_k)}.$$

Then the generalized controllability matrix (along the sequence $\{\xi_i ; 0 \leq i \leq k\}$), $C^k_x = C^k_x(\xi_0, \ldots, \xi_k)$, is defined by

$$C^k_x \equiv [A_k \ldots A_1 B_0 \mid A_k \ldots A_2 B_1 \mid \ldots \mid A_k B_{k-1} \mid B_k].$$

Remark 1.

As pointed out in Meyn and Caines (1988), the term generalized controllability matrix is justified by the fact that, if the dynamical system above is simply a linear system with $W = \mathbb{R}^n$ and $M = \mathbb{R}^d$ (i.e., $f(x, \xi) = Ax + B\xi$, with both A and $B \in gl(d, \mathbb{R})$), $C_x^k = [A^k B \mid A^{k-1} B \mid \ldots \mid A B \mid B]$, is then simply the usual controllability matrix at time k.

To use generalized controllability matrices, we first need the following lemma, which is an extension of Lemma 2.3 in Meyn and Caines (1988).

Lemma 1.

Let M_1^d, M_2^m, and M_3^d be C^∞ manifolds and let m_{M_1}, m_{M_2}, and m_{M_3} denote the volume elements on M_1, M_2, and M_3, respectively. Let $O_1 \subset M_1$, $U_1 \subset M_2$, and $V_1 \subset M_3$ be open. Suppose that $G : O_1 \times U_1 \times V_1 \to M_1$ is C^1 and that the matrix $\frac{\partial G}{\partial z}$ is full rank at some $(x_0, y_0, z_0) \in O_1 \times U_1 \times V_1$. Then we have:

a) There exists an open set $O \times U \times V \subset O_1 \times U_1 \times V_1$ containing (x_0, y_0, z_0) such that the measure $\nu(x, .)$ defined for $A \in \mathcal{B}(M_1)$ and $x \in O$ by

$$\nu(x, A) = \int_U \int_V 1_{[G(x,y,z) \in A]}(x, y, z)\, m_{M_2}(dy)\, m_{M_3}(dz)$$

is equivalent to m_{M_1} on an open set $R_x \subset M_1$, where 1_A denotes the indicator function of the set A. (R_x also depends on U and V.)

b) There exist $c > 0$ and open sets $P \subset M_1$ and $T \subset M_1$, containing x_0 and $G(x_0, y_0, z_0)$ respectively, such that, for all $x \in M_1$,

$$\nu(x, .) \geq c\, 1_P(x)\, m_{M_1}(T \cap .).$$

(c, P, and T depend on the sets U and V from part (a).)

Proof.

a) Let (W_1, φ_1), (W_2, φ_2), (W_3, γ_3), and (W', ψ) be coordinate neighborhoods of x_0, y_0, z_0, and $G(x_0, y_0, z_0)$ respectively.

Define $\varphi : W_1 \times W_2 \times W_3 \longrightarrow \mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^d$ by

$$\varphi(x, y, z) = (\varphi_1(x), \varphi_2(y), \varphi_3(z)),$$

so that $(W_1 \times W_2 \times W_3, \varphi) \equiv (W, \varphi)$ is a coordinate neighborhood of (x_0, y_0, z_0).

Then $\varphi(W \cap G^{-1}(W') \cap [O_1 \times U_1 \times V_1])$ is an open set in $\mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^d$ containing some further open set of product form $O_1^j \times U_1^j \times V_1^j$ (O_1^j, U_1^j, and V_1^j open) with $\varphi(x_0, y_0, z_0) \equiv (u_0, v_0, w_0) \in O_1^j \times U_1^j \times V_1^j$.

Moreover, the mapping $G' : O_1^j \times U_1^j \times V_1^j \longrightarrow \psi(W')$ defined by $G' = \psi \circ G \circ \varphi^{-1}$ is C^1 and satisfies the condition that $\frac{\partial G'}{\partial z}$ is full rank at the point $(u_0, v_0, w_0) \in O_1^j \times U_1^j \times V_1^j$.

Hence, the conditions of Lemma 2.3 in Meyn and Caines (1988) are satisfied for the open set $O_1^j \times U_1^j \times V_1^j$ and the mapping G'. It follows that there exists an open set $O' \times U' \times V'$ containing (u_0, v_0, w_0) such that the measure $\tilde{\nu}(u, .)$ defined for $u \in O'$ and $B \in \mathcal{B}(\mathbb{R}^d)$ by

$$\tilde{\nu}(u, B) = \int_{U'} \int_{V'} 1_{[G'(u, v, w) \in B]}(u, v, w)\, dw\, dv$$

is equivalent to the Lebesgue measure on an open set $R_u' \subset \psi(W') \subset \mathbb{R}^d$. (The proof of this result indicates that R_u' also depends on O', U', and V'.)

Let $O \times U \times V \equiv \varphi^{-1}(O' \times U' \times V')$ and define $\nu(x, .)$ for $x \in O$ and $A \in \mathcal{B}(M_1)$ by $\nu(x, A) = \tilde{\nu}(\varphi_1(x), \psi(A \cap W'))$. By Corollary 6.1.14 in Boothby (1986), for

any fixed $x \in O$, $\nu(x,.)$ is equivalent to m_{M_1} on $\psi^{-1}(R'_p)$, where $p = \varphi_1(x)$.

b) Using the notation from part (a), Lemma 2.3 in Meyn and Caines (1988) shows that there exist $c' > 0$ and open rectangles P' and T' containing u_o and G' (u_o, v_o, w_o) respectively such that, for all $u \in \mathbb{R}^d$,

$$\tilde{\nu}(u,.) \geq c' 1_{P'}(u) m_{\mathbb{R}^d}(T' \cap .),$$

where $m_{\mathbb{R}^d}$ represents the Lebesgue measure on $\mathbb{R}^d$. (The proof of this result also indicates that c', P', and T' depend on the sets O', U', and V' from part (a).) Then, as in part (a), for all $B \in \mathcal{B}(M_1)$ and all $x \in M_1$,

$$\begin{aligned}
\nu(x,B) &= \tilde{\nu}(\varphi_1(x), \psi(W' \cap B)) \\
&\geq c' 1_{P'}(\varphi_1(x)) m_{\mathbb{R}^d}(T' \cap \psi(W' \cap B)) \\
&\geq c' 1_P(x) \frac{m_{\mathbb{R}^d}(T' \cap \psi(W' \cap B))}{m_{\mathbb{R}^d}(T')} m_{\mathbb{R}^d}(T') \frac{m_{M_1}(T \cap B)}{m_{M_1}(T)} \\
&\geq c\, 1_P(x)\, m_{M_1}(T \cap B),
\end{aligned}$$

with $P = \varphi^{-1}(P')$, $T = \psi^{-1}(T') \cap W'$, and $c = c' \frac{m_{\mathbb{R}^d}(T')}{m_{M_1}(T)}$ (noting that $m_{M_1}(T)$ is finite and nonzero since ψ^{-1} is continuous and T' is a bounded open set).

This completes the proof. ∎

We are now ready to state the following result.

<u>Theorem 1.</u>

Assume that $\{\xi_n ; n \geq 0\}$ is a Markov chain on a C^∞ manifold W^k and that x_o is random variable taking values in a C^∞ manifold M^d. Also assume that x_o is

independent of the process $\{\xi_n\}$, and that $\{\xi_n\}$ is ergodic with invariant probability measure π_ξ satisfying $\pi_\xi(\text{int } Q) = \pi_\xi(Q)$ with $Q \equiv \text{supp}(\pi_\xi)$. Assume moreover that the absolutely continuous component (with respect to the volume element in W) of the one step transition probability $\mu(\xi, .)$ of the $\{\xi_n\}$ process, $g(\xi, \xi')$, is, for each $\xi \in \text{int } Q$, strictly positive on int Q and lower semi-continuous in both of its arguments.

Then the following holds:

a) The stochastic dynamical system $x_{n+1} = f(x_n, \xi_n)$, with $f: M \times W \longrightarrow M$ being a C^1 map, satisfies Assumption 7 if and only if, for all initial conditions $(x, \xi) \in M \times \text{int } Q$, there exists a time $t \geq 1$ ($t \in \mathbb{N}$) and $\underset{\sim}{\xi}_t \equiv (\xi_1, \ldots, \xi_t) \in [\text{int } Q]^t$ (where $[\text{int } Q]^t$ represents the t-fold cartesian product of int Q) such that

$$\text{rank } C_x^t(\underset{\sim}{\xi}_t) = d$$

b) If $\text{rank } C_x^t(\underset{\sim}{\xi}_t) = d$ for some $\underset{\sim}{\xi}_t \in [\text{int } Q]^t$, then there exist $c > 0$ and open sets $U \subset M$ and $V \subset M$ containing x and $S_{x\,\xi}^{t+1}$ respectively, and an open set $R \subset \text{int } Q$, such that, for all $B \in \mathcal{B}(M)$ and all $(x, \xi) \in U \times R$,

$$P(S_{x\,\xi}^{t+1} \in B \mid (x_0, \xi_0) = (x, \xi)) \geq c\, m_M(B \cap V),$$

where m_M denotes the volume element on M and

$$\begin{aligned} &P(S_{x\,\xi}^{t+1} \in B \mid (x_0, \xi_0) = (x, \xi)) \\ &\quad = P([S_{x\,\xi}^{t+1}, \xi_t] \in B \times W \mid (x_0, \xi_0) = (x, \xi)) \\ &\quad = P([S_{x\,\xi}^{t+1}, \xi_t] \in B \times \text{int } Q \mid (x_0, \xi_0) = (x, \xi)), \end{aligned}$$

the last equality holding because of the invariance of int Q (see the discussion of

Assumption 6).

Proof.

a) The basic argument can be found for the most part in Meyn and Caines (1988, Theorem 2.1 (ii)).

The necessity of the rank condition follows from Sard's Theorem (see Meyn and Caines (1988, proof of Theorem 2.1), or Jacubczyk and Sontag (1988, proof of Proposition 2.3)).

To prove sufficiency, first note that we may always assume $t > d$ since, if rank $C_y^t(\underset{\sim}{\xi}_t) = d$, then rank $C_y^{t'}(\underset{\sim}{\xi}_t) = d$ for all $t' \geq t$. Let $\xi \in \text{int } Q$ be fixed, $B \in \mathcal{B}(M)$, and suppose that the rank condition is satisfied for some $x \in M$ and some $\underset{\sim}{\xi}_t^0 = (\xi_1^0, \ldots, \xi_t^0) \in [\text{int } Q]^t$. Then,

$$P(S_{x\,\xi}^{t+1} \in B \mid x_0 = x, \xi_0 = \xi)$$

$$= \int_W \cdots \int_W 1_{[S_{x\,\xi}^{t+1} \in B]}(\xi_1, \ldots, \xi_t)\, P(\xi_1 \in d\xi_1, \ldots, \xi_t \in d\xi_t \mid \xi_0 = \xi)$$

$$= \int_W \cdots \int_W 1_{[S_{x\,\xi}^{t+1} \in B]}(\xi_1, \ldots, \xi_t)\, g(\xi_{t-1}, \xi_t) \cdots g(\xi, \xi_1)\, m_W(d\xi_1 \times \ldots \times d\xi_t).$$

Now, since $g(\xi, .)$, $\xi \in \text{int } Q$ fixed, is lower semi–continuous, we can find an open set $V_1 \subset \text{int } Q$ and containing ξ_1^0 such that $g(\xi, \xi_1) \geq p_1 > 0$ for $\xi_1 \in V_1$. Similarly, for any fixed $\xi_1 \in V_1 \subset \text{int } Q$, we can find an open set $V_2 \subset \text{int } Q$ and containing ξ_2^0 such that $g(\xi_1, \xi_2) \geq p_2 > 0$ for $\xi_2 \in V_2$. Lower semi–continuity in the ξ_1 variable then implies that there is in fact an entire open set U_1 containing ξ_1^0 and such that $g(\xi_1, \xi_2) \geq p_2$ for $\xi_1 \in U_1$, $\xi_2 \in V_2$. This implies that, for the pair $(\xi_1, \xi_2) \in O_1 \times V_2$, $O_1 \equiv V_1 \cap U_1$, we have $g(\xi, \xi_1)\, g(\xi_1, \xi_2) \geq p_1 p_2 > 0$.

Repeating the above argument for $g(\xi_2, \xi_3)$ up to $g(\xi_{t-1}, \xi_t)$ shows that, whenever $\underset{\sim}{\xi}_t \in H(\underset{\sim}{\xi}^0_t) \equiv O_1 \times \ldots \times O_{t-1} \times V_t \subset [\text{int } Q]^t$ we have

$$g(\xi_{t-1}, \xi_t) \ldots g(\xi, \xi_1) \geq p_1 \ldots p_t \equiv p > 0.$$

(Note that both p and $H(\underset{\sim}{\xi}^0_t)$ depend on ξ but that lower semi–continuity ensures that a single choice for p and $H(\underset{\sim}{\xi}^0_t)$ can be made for all $\xi \in R \subset W$, R open.)

Hence, we get, for all $\xi \in R$,

$$P(S_x^{t+1}{}_{\xi} \in B \mid x_0 = x, \xi_0 = \xi)$$

$$\geq p \int_{H(\underset{\sim}{\xi}^0_t)} 1_{[S_x^{t+1}{}_{\xi} \in B]}(\xi_1, \ldots, \xi_t)\, m_W(d\xi_1 \times \ldots \times d\xi_t).$$

But then, defining the mapping $G = G_\xi : M^d \times (\text{int } Q)^{t-d} \times (\text{int } Q)^d \longrightarrow M^d$ by $G_\xi(x, (\xi_t, \ldots, \xi_{d+1}), (\xi_d, \ldots, \xi_1)) = S_x^{t+1}{}_{\xi}$, the condition rank $C_x^t(\underset{\sim}{\xi}^0_t) = d$ does imply that there must be integers $\{i_1, \ldots, i_d\}$ such that, defining the stochastic orbit of the point $x \in M$ by $S_{x\,\xi} \equiv \bigcup \{S_{x\,\xi}^t ; t \geq 1\}$,

$$\det \left[\frac{\partial S_{x\,\xi}}{\partial \xi_{i_d}} \Big|, \ldots, \Big| \frac{\partial S_{x\,\xi}}{\partial \xi_{i_1}} \right]_{(x, \xi, \underset{\sim}{\xi}^0_t)} \neq 0$$

(see the proof of the next proposition). This allows us to use the result of Lemma 1 (a) to claim that, for all $\xi \in R \subset \text{int } Q$, there must exists an open set $T \subset M$ (which will depend on x and $\underset{\sim}{\xi}^0_t$) such that, for all $B \in \mathcal{B}(M)$ and with $H'(\underset{\sim}{\xi}^d_t) \equiv O_{d+1} \times \ldots \times O_{t-1} \times V_t$ and $H(\underset{\sim}{\xi}_d) \equiv O_1 \times \ldots \times O_d$,

$$\int_{H(\underset{\sim}{\xi}_t^o)} 1_{[S_x^{t+1}\,\xi \in B]}(\xi_1, \ldots, \xi_t)\, m_W(d\xi_1 \times \ldots \times d\xi_t)$$

$$= \int_{H(\underset{\sim}{\xi}_d)} \int_{H'(\underset{\sim}{\xi}_t^d)} 1_{[S_x^{t+1}\,\xi \in B]}(\xi_1, \ldots, \xi_t)\, m_W(d\xi_1 \times \ldots \times d\xi_t)$$

is equivalent to $m_M(T \cap B)$.

This proves weak stochastic controllability (in the M component).

b) By the reasoning in part (a) and using Lemma 1 (b), we also conclude that, for some $c > 0$ and for some open sets U and V (both depending on x and $\underset{\sim}{\xi}_t^o$) containing x and $S_x^{t+1}\,\xi$ respectively, we have, for all $(x, \xi) \in M \times R$,

$$P(S_x^{t+1}\,\xi \in B \mid x_o = x, \xi_o = \xi) \geq c\, 1_U(x)\, m_M(B \cap V),$$

which proves the second statement of Theorem 1. ∎

Remark 2.

One case of special interest to us in Section 5 is the canonical action of a subsemigroup of $Gl(d, \mathbb{R})$ (arising from the system $s_{n+1} = A(\xi_n)\, s_n \,\|A(\xi_n)\, s_n\|^{-1}$) on the projective space $\mathbb{P}^{d-1}$. We could apply the above result directly to this setup but, working with this nonlinear action (which, in our case, will come from the projection of the linear system $x_{n+1} = A(\xi_n)\, x_n$ in $\mathbb{R}_0^d$) is more difficult than working with the unprojected (linear) system.

Now, since $\mathbb{P}^{d-1}$ (or $\mathbb{S}^{d-1}$) are (lower dimensional) subspaces of $\mathbb{R}_0^d$ which are left invariant under $s_{n+1} = A(\xi_n)\, s_n \,\|A(\xi_n)\, s_n\|^{-1}$, it is enough to show that the rank condition of Theorem 1 is satisfied for the unprojected system $x_{n+1} = A(\xi_n)\, x_n$,

i.e., for the mappings $A(\xi_n)x$. Indeed, if the generalized controllability matrix of the system in $\mathbb{R}_0^d$ satisfies the rank condition of Theorem 1, and hence we have weak stochastic controllability for the unprojected system, we also automatically have weak stochastic controllability for the same system projected onto $\mathbb{P}^{d-1}$ (or $\mathbb{S}^{d-1}$) since the projection onto $\mathbb{P}^{d-1}$ (or $\mathbb{S}^{d-1}$) of sets with nonzero Lebesgue measure in $\mathbb{R}_0^d$ gives sets of nonzero measure with respect to the volume element on $\mathbb{P}^{d-1}$ (or $\mathbb{S}^{d-1}$). This fact could (but will not) be used in Section 6 to verify the assumption of weak stochastic stability for the angular behavior of a discretized version of the linear oscillator with damping and restoring force on $\mathbb{R}_0^2$.

We are now ready to complete our investigation of Assumption 7 and its relationship to the condition int $Sx \neq \phi$ for all $x \in M$. In Meyn and Caines (1988), the authors state that, under their setup, both the rank condition of Theorem 1 and the condition int $Sx \neq \phi$ for all $x \in M$ are equivalent to weak stochastic stability (footnote p. 11). The next proposition proves this assertion and shows that same holds true under our assumptions.

Proposition 1.

Consider the deterministic control system on M^d

$$x_{n+1} = f(x_n, u_n),\ u_n \in U,\ U \text{ open in } W,\ f: M \times W \longrightarrow M \text{ is } C^1.$$

Then, int $Sx \neq \phi$ for all $x \in M$ if and only if there exists $n \in \mathbb{N}$ (which depends on x) such that rank $C_x^n = d$.

Proof.

By Proposition 2.3 in Jacubczyk and Sontag (1988), int $Sx \neq \phi$ for all $x \in M$ if and only if

$$\sup\left\{\text{rank}\left[\frac{\partial}{\partial(u_0,\ldots,u_k)} f_{u_k\ldots u_0}(x)\,;(u_0,\ldots,u_k)\in U^{k+1}, k\geq 0\right]\right\} = d.$$

(Note that Jacubczyk and Sontag use $W = \mathbb{R}^k$, but a rank condition at a point is a local property and hence, their result immediately applies to the manifold W.)

Write $y_k(x) \equiv \frac{\partial}{\partial(\zeta_0,\ldots,\zeta_k)} f_{\zeta_k\ldots\zeta_0}(x)\Big|_{(u_0,\ldots,u_k)}$ and note that, for $k \geq 1$,

$$A_k = \frac{\partial f}{\partial x}\Big|_{(f_{u_{k-1}\ldots u_0}(x),\, u_k)},$$

$$B_k = \frac{\partial f}{\partial u}\Big|_{(f_{u_{k-1}\ldots u_0}(x),\, u_k)}, \text{ and } B_0 = \frac{\partial f}{\partial u}\Big|_{(x,\,u_0)} = \frac{\partial}{\partial u} f(x,u)\Big|_{u_0}.$$

Then, $y_0(x) = \frac{\partial}{\partial \zeta_0} f_{\zeta_0}(x)\Big|_{u_0} = \frac{\partial}{\partial \zeta_0} f(x,\zeta_0)\Big|_{u_0} = B_0.$

Similarly, $y_1(x) = \frac{\partial}{\partial(\zeta_0,\zeta_1)} f_{\zeta_1\zeta_0}(x)\Big|_{(u_0,u_1)} = \frac{\partial}{\partial(\zeta_0,\zeta_1)} f(f(x,\zeta_0),\zeta_1)\Big|_{(u_0,u_1)}$

$$= \frac{\partial}{\partial \zeta_0} f(f(x,\zeta_0),\zeta_1)\Big|_{(u_0,u_1)} d\zeta_0 + \frac{\partial}{\partial \zeta_1} f(f(x,\zeta_0),\zeta_1)\Big|_{(u_0,u_1)} d\zeta_1$$

$$= \left[\frac{\partial f}{\partial x}\Big|_{(f(x,u_0),u_1)} \cdot \frac{\partial}{\partial \zeta_0} f(x,\zeta_0)\Big|_{u_0}\right] d\zeta_0 + \frac{\partial f}{\partial u_1}\Big|_{(f(x,u_0),u_1)} d\zeta_1$$

$$= A_1 B_0\, d\zeta_0 + B_1\, d\zeta_1 = \left[A_1 B_0 \mid B_1\right]\begin{bmatrix} d\zeta_0 \\ d\zeta_1 \end{bmatrix}.$$

In general, if we have

$$y_{k-1}(x) = \left[A_{k-1}\cdots A_1 B_0 \mid A_{k-1}\cdots A_2 B_1 \mid \ldots \mid A_{k-1} B_{k-2} \mid B_{k-1}\right]\begin{bmatrix} d\zeta_0 \\ \vdots \\ d\zeta_{k-1} \end{bmatrix},$$

we get,

$$y_k(x) = \frac{\partial}{\partial(\zeta_0, \dots, \zeta_k)} f_{\zeta_k \dots \zeta_0}(x)\Big|_{(u_0, \dots, u_k)}$$

$$= \frac{\partial}{\partial(\zeta_0, \dots, \zeta_k)} f\left(f_{\zeta_{k-1} \dots \zeta_0}(x), \zeta_k\right)\Big|_{(u_0, \dots, u_k)}$$

$$= \frac{\partial}{\partial \zeta_0} f\left(f_{\zeta_{k-1} \dots \zeta_0}(x), \zeta_k\right)\Big|_{(u_0, \dots, u_k)} d\zeta_0$$

$$+ \dots + \frac{\partial}{\partial \zeta_{k-1}} f\left(f_{\zeta_{k-1} \dots \zeta_0}(x), \zeta_k\right)\Big|_{(u_0, \dots, u_k)} d\zeta_{k-1}$$

$$+ \frac{\partial}{\partial \zeta_k} f\left(f_{\zeta_{k-1} \dots \zeta_0}(x), \zeta_k\right)\Big|_{(u_0, \dots, u_k)} d\zeta_k$$

$$= \left[\frac{\partial f}{\partial x}\Big|_{\left(f_{u_{k-1} \dots u_0}(x), u_k\right)} \cdot \frac{\partial}{\partial \zeta_0} f_{\zeta_{k-1} \dots \zeta_0}(x)\Big|_{(u_0, \dots, u_{k-1})}\right] d\zeta_0$$

$$+ \dots + \left[\frac{\partial f}{\partial x}\Big|_{\left(f_{u_{k-1} \dots u_0}(x), u_k\right)} \cdot \frac{\partial f}{\partial \zeta_{k-1}}\Big|_{\left(f_{u_{k-2} \dots u_0}(x), u_{k-1}\right)}\right] d\zeta_{k-1}$$

$$+ \frac{\partial f}{\partial \zeta_k}\Big|_{\left(f_{u_{k-1} \dots u_0}(x), u_k\right)} d\zeta_k$$

$$= \left[\frac{\partial f}{\partial x}\Big|_{\left(f_{u_{k-1} \dots u_0}(x), u_k\right)} A_{k-1} \cdots A_1 B_0\right] d\zeta_0$$

$$+ \dots + \left[\frac{\partial f}{\partial x}\Big|_{\left(f_{u_{k-1} \dots u_0}(x), u_k\right)} \cdot B_{k-1}\right] d\zeta_{k-1}$$

$$+ \frac{\partial f}{\partial \zeta_k}\Big|_{\left(f_{u_{k-1} \dots u_0}(x), u_k\right)} d\zeta_k$$

$$= \left[A_k \, A_{k-1} \cdots A_1 \, B_0 \mid A_k \, A_{k-1} \cdots A_2 \, B_1 \mid \cdots \mid A_k \, B_{k-1} \mid B_k \right] \begin{bmatrix} d\zeta_0 \\ \vdots \\ d\zeta_k \end{bmatrix}$$

$$= C_x^k \begin{bmatrix} d\zeta_0 \\ \vdots \\ d\zeta_k \end{bmatrix}.$$

From this it follows that, for all $x \in M$ and for some sequence $\{u_0, \ldots, u_k\}$, $k \geq 0$,

$$\text{rank } C_x^k = d,$$

if and only if, for all $x \in M$,

$$\sup \left\{ \text{rank} \left[\frac{\partial}{\partial (u_0, \ldots, u_k)} f_{u_k \ldots u_0}(x) \, ; \, (u_0, \ldots, u_k) \in U^{k+1}, k \geq 0 \right] \right\} = d,$$

and this completes the proof. ■

Hence the discrete time case resembles the continuous time case in the sense that, under the lower semi–continuity assumption made on the densities of the one–step transition probabilities $\mu(\xi, .)$ (Assumption 6), Assumptions 7 is implied by the condition int $Sx \neq \phi$ for all $x \in M$, which holds true if and only if some Lie algebra of vector fields has full dimension (see Theorem 2.3.2). In other words, as in the continuous time case, some Lie algebra condition turns out to be an underlying but crucial assumption for the approach described in this work.

Assumption 8 solely depends on the dynamics of the associated control system and thus cannot be implied by our assumptions on the process $\{\xi_n\}$, or by a condition of the type int $Sx \neq \phi$ for all $x \in M$ (which, as discussed above, is related to assumptions on the process $\{\xi_n\}$). In fact, in general, a nonlinear control system

may have several invariant control sets. Nevertheless, for nonlinear control systems on $\mathbb{P}^{d-1}$ and arising from the projection of linear systems in $\mathbb{R}^d_0$ onto $\mathbb{P}^{d-1}$, we know that int $Ss \neq \phi$ for all $s \in \mathbb{P}^{d-1}$ does imply the existence of a unique maximal invariant control set on $\mathbb{P}^{d-1}$ (see Theorem 3.3.1). This setup will be used in Sections 5 and 6.

4.2. Ergodic Behavior of Stochastic Difference Equations

Consider the stochastic difference equation described in Subsection 4.1, i.e.,

$$x_{n+1} = f(x_n, \xi_n), \quad n \geq 0, \; x_0 = x_0(\omega),$$

and further assume that the Assumptions 1 through 8, stated and discussed in Subsection 4.1, hold.

The cornerstone of our study of such a dynamical system reposes on the existence of a unique invariant probability measure for the pair process $\{(x_n, \xi_n) ; n \in \mathbb{N}\}$. The aim of this section is to show that, under the given assumptions, the Markov chain $\{(x_n, \xi_n) ; n \in \mathbb{N}\}$ is ergodic, i.e., admits such a unique invariant probability measure, which will be denoted by π.

In order to obtain such a result, we refer to Subsection 2.8 and adapt Theorem 2.8.1 to our setup. We will first discuss an argument which will be extremely useful in several of the following proofs. Next, we will prove several results, and then verify Conditions (A) and (B) from Subsection 2.8.

Decoupling of the $\{(x_n, \xi_n)\}$ process.

Consider, under our setup, the Markov process $\{(x_n, \xi_n)\}$. As in Subsection 4.1, denote $f_{\xi_{n-1}\cdots\xi_0}(x)$ by $S^n_{x\,\xi} \in M$ ($S^0_{x\,\xi} \equiv x$) and write $S_{x\,\xi} \equiv \bigcup_{k=1}^{\infty} S^k_{x\,\xi} \subset M$ for the (stochastic) orbit up to time n of the point $x \in M$, starting at the initial value $\xi_0 = \xi \in W$.

Hence, starting at some initial value $(x_0, \xi_0) = (x, \xi)$, $(x_n, \xi_n) = (S^n_{x\,\xi}, \xi_n)$ can be written as $(f_{\xi_{n-1}\cdots\xi_0}(x), \xi_n)$. This shows that the value of x_n only depends on $(\xi_{n-1}, \ldots, \xi_0)$ and that constraints on the value of x_n do not affect ξ_n (by the Markov property, ξ_n still depends on ξ_{n-1}). Therefore, if $\xi \in$ int Q with $Q \equiv$ supp (π_ξ), Assumption 5 and the invariance of int Q (see the discussion of Assumption 6 in Subsection 4.1) guarantee that, for any values $x' \in M$ and $\xi' \in$ int Q of x_n and ξ_{n-1}, $P(\xi_n \in B \mid \xi_{n-1} = \xi') > 0$ for all $B \in \mathcal{B}(W)$ such that $\pi_\xi(B \cap Q) > 0$. This argument will be referred to as the "Decoupling Argument" and will be used in several of the ensuing proofs.

Lemma 1

Let $\{y_n ; n \in \mathbb{N}\}$ be a Markov chain valued on some topological space E.

Assume that for some sets A, B, and $D \in \mathcal{B}(E)$, $L(x, B) > 0$ for all $x \in A$ and $L(y, D) > 0$ for all $y \in B$.

Then $L(x, D) > 0$ for all $x \in A$.

Proof.

By Remark 2.7.2, we have $G(x, B) > 0$ for all $x \in A$ and $G(y, D) > 0$ for all $y \in B$ (both $G(x, B)$ and $G(y, D)$ possibly infinite). Hence, for all $x \in A$, we have

$0 < \sum_{n=1}^{\infty} \left[\frac{1}{2}\right]^n \mu(n, x, B) < 1$, where $\mu(n, x, B)$ is used to denote the n^{th} step transition probability from x to B. Therefore, for all $x \in A$,

$$0 < \sum_{n=1}^{\infty} \left[\frac{1}{2}\right]^n \int_B G(y, D)\, \mu(n, x, dy)$$

$$\leq G(x, D) \sum_{n=1}^{\infty} \left[\frac{1}{2}\right]^n = G(x, D).$$

But, again, for any $x \in A$, $G(x, D) > 0$ if and only if $L(x, D) > 0$ (Remark 2.7.2), and this completes the proof.

■

Lemma 2 ("Tube method" (see Arnold and Kliemann (1983))

Let $k \in \mathbb{N}$ be arbitrary. Then, for all $\epsilon > 0$, all initial values $\xi \in \text{int } Q$, and all sequences $\{u_1, \ldots, u_k\} \subset [\text{int } Q]^k$,

$$P\left[\max_{1 \leq i \leq k} \rho(\xi_i, u_i) < \epsilon \mid \xi_0 = \xi\right] > 0,$$

where ρ denotes the metric on the manifold W (see Theorem 2.3.1).

Proof.

For each u_i, $1 \leq i \leq k$, the condition $\rho(\xi_i, u_i) < \epsilon$ describes an open set $\kappa_i \subset \text{int } Q$. So,

$$P\left[\max_{1 \leq i \leq k} \rho(\xi_i, u_i) < \epsilon \mid \xi_0 = \xi\right] = P(\xi_1 \in \kappa_1, \ldots, \xi_k \in \kappa_k \mid \xi_0 = \xi).$$

For each $1 \leq i \leq k$, let $D_i \subset \kappa_i$ be a closed set such that $\pi_\xi(D_i) > 0$. For any $B \in \mathcal{B}(W)$, let $c(i, B) \equiv \inf\{\mu(\xi, B)\,;\, \xi \in D_i\}$. Then, by Assumptions 5 and 6, for $1 \leq i \leq k$, $c(i, \kappa_{i+1}) > 0$ and $\mu(\xi, D_i) > 0$ for all $\xi \in \text{int } Q$. Hence,

$$P(\xi_1 \in \kappa_1, \ldots, \xi_k \in \kappa_k \mid \xi_0 = \xi)$$

$$= \int_{\kappa_1 \times \ldots \times \kappa_{k-1}} \mu(\xi_{k-1}, \kappa_k)\, \mu(\xi_{k-2}, d\xi_{k-1}) \ldots \mu(\xi_1, d\xi_2)\, \mu(\xi, d\xi_1)$$

$$\geq c(k\text{-}1, \kappa_k) \int_{D_1 \times \ldots \times D_{k-2}} \mu(\xi_{k-2}, D_{k-1}) \ldots \mu(\xi_1, d\xi_2)\, \mu(\xi, d\xi_1)$$

$$\geq c(k\text{-}1, \kappa_k) \ldots c(1, D_2)\, \mu(\xi, D_1) > 0.$$

This completes the proof. ■

Using the above lemma, we are now ready to prove a basic result, which will be used repeatedly in the remainder of this subsection.

Proposition 1.

For any open sets $O \subset C$ (the unique maximal invariant control set for the associated dynamical control system) and any set $U \in Q$ satisfying $\pi_\xi(U) > 0$, we have

$$L((x, \xi), O \times U) = P\left[\bigcup_{n=1}^{\infty} [S^n_{x\,\xi} \in O, \xi_n \in U] \mid (x_0, \xi_0) = (x, \xi)\right] > 0$$

for all pairs $(x, \xi) \in C \times W$.

Moreover, if M is compact, the above holds for all pairs $(x, \xi) \in M \times W$.

Proof

First assume that $(x, \xi) \in C \times \text{int } Q$. Since C is a maximal invariant control set, it is clear that there exists a deterministic path from $f(x, \xi)$ to any open set $O \subset C$.

Denote this path by

$$\{f_{u_k \ldots u_0}(f(x, \xi)) ; 0 \leq k \leq n\},\ f_{u_n \ldots u_0}(f(x, \xi)) \in O.$$

Write $B[x, r]$ for an open ball of radius $r > 0$ centered at x.

By continuity of f, there exists a sequence of products of open balls

$$\left\{W_{\delta_k} \times V_{\epsilon_k} ; 0 \le k \le n\text{-}1\right\}$$

$$\equiv \left\{B\left[f_{u_k \ldots u_0}(f(x,\xi)), \delta_k\right] \times B\left[u_{k+1}, \epsilon_k\right] ; 0 \le k \le n\text{-}1\right\} \subset C \times \text{int } Q,$$

$Q = \text{supp}(\pi_\xi)$, and, with $W_{\delta_n} \times V_{\epsilon_n} \subset O \times U$, such that $(x, u) \in W_{\delta_{k-1}} \times V_{\epsilon_{k-1}}$ does imply $f(x, u) \in W_{\delta_k}$. In other words, $f_{u_{k-1} \ldots u_0}(f(x, \xi)) \in W_{\delta_k}$ for all k–tuples $(u_{k-1}, \ldots, u_0) \in U_{\epsilon_{k-1}} \times \ldots \times U_{\epsilon_0}$ (note that ϵ_k depends on δ_k and δ_{k+1}, $0 \le k \le n\text{-}1$). But, by Lemma 2,

$$P(\xi_1 \in V_{\epsilon_0}, \ldots, \xi_{n+1} \in V_{\epsilon_n} \mid \xi_0 = \xi)$$

$$> P\left[\max_{1 \le k \le n+1} \rho(\xi_k, u_k) < \min_{1 \le k \le n+1} \epsilon_k \mid \xi_0 = \xi\right] > 0,$$

from which it follows that, for all $(x, \xi) \in C \times \text{int } Q$ (using the invariance of int Q),

$$L((x,\xi), O \times U) = P\left[\bigcup_{n=1}^{\infty} [S_{x\,\xi}^n \in O, \xi_n \in U] \mid (x_0, \xi_0) = (x, \xi)\right] > 0.$$

If, more generally, $(x, \xi) \in C \times W$, note that, by weak recurrence of the $\{\xi_n\}$ process, $G(\xi, \text{int } Q) = \infty$ for all $\xi \in W$. Hence, for all $(x, \xi) \in C \times W$, we have $L((x, \xi), C \times \text{int } Q) > 0$. Therefore, by Lemma 1, $L((x, \xi), O \times U) > 0$ is true for all $(x, \xi) \in C \times W$.

Finally, if M is compact, by Proposition 3.1.12, $\phi \neq C \subset \overline{Sx}$. Therefore, int $C \neq \phi$ implies $C \cap Sx \neq \phi$ and we have, for all $(x, \xi) \in M \times \text{int } Q$, a deterministic path from $f(x, \xi)$ to any open set $O \subset C$. The result then follows from the same argument as above.

■

Remark 1.

1) We will show later on that the statements of Proposition 1 can be "strengthened" (see Theorem 3), but this strengthening will first require other results (and assumptions).

2) In fact, the proof of Proposition 1 shows that, with $O \subset C$, O open, and $U \subset \text{int } Q$ with $\pi_\xi(U) > 0$, we have $L((x, \xi), O \times U) > 0$ for all $(x, \xi) \in M \times W$ such that $L((x, \xi), C \times \text{int } Q)$ is not zero, i.e., whenever one can guarantee (e.g., from more specific information about the dynamical system under scrutiny) a positive probability of entering the set C, starting with the initial value (x, ξ). The compactness of M is simply a general assumption ensuring that this holds true for all $(x, \xi) \in M \times W$.

Remark 2.

The uniqueness of the maximal invariant control set C and Proposition 1 easily imply that the M component of any invariant set for the pair process $\{(x_n, \xi_n)\}$ must be dense in C.

Now, using our assumptions and with the help of Proposition 1, we are able to proceed with the proof of the existence of a unique invariant measure for the Markov chain $\{(x_n, \xi_n)\}$. First define a measure m_C on $(M, \mathcal{B}(M))$ by

$$m_C(B) = m_M(B \cap C),$$

where m_M denotes the volume element on the manifold M. Note that m_C is nontrivial since C has a nonempty interior. Also, if M is compact, m_C is a finite measure and hence can be standardized to a probability measure.

Lemma 3.

Every invariant set $I \subset M \times W$ under the Markov chain $\{(x_n, \xi_n)\}$ contains a set of the form $\Gamma \equiv \{(x, \xi) \in M \times W : x \in \Gamma_M, \xi \in \Gamma_W(x)\}$ with $\Gamma_M \subset C$ dense in C and, for every $x \in \Gamma_M$, $\Gamma_W(x)$ satisfying $\pi_\xi(\Gamma_W(x)) = 1$.

Proof.

Under our setup, the set $C \times Q \subset M \times W$ is certainly an invariant set for the pair process $\{(x_n, \xi_n)\}$ (because C is C–invariant for the associated dynamical control system with controls valued in int Q with $\pi_\xi(\text{int } Q) = 1$ (Assumption 5)). Let $I \subset M \times W$ be any other invariant set for $\{(x_n, \xi_n)\}$. By Proposition 1, we clearly have that $I \cap (C \times \text{int } Q) \neq \phi$. Pick a point $(x, \xi) \in I \cap (C \times \text{int } Q)$. By Remark 2, the stochastic orbit of (x, ξ), $S_{x\,\xi}$, will (with positive probability) be dense in C. By the decoupling of (x_n, ξ_n) and since, by the invariance of int Q, ξ_n is valued in int Q for all $n \in \mathbb{N}$, whatever the value of x_n, ξ_n can be valued, with positive probability, in any subset of W having positive π_ξ measure. This implies that some set of the form

$$\Gamma \equiv \{(x, \xi) \in M \times W : x \in \Gamma_M, \xi \in \Gamma_W(x)\}$$

is included in I, with $\Gamma_M \subset C$ dense in C and, for all $x \in \Gamma_M$, $\Gamma_W(x) \subset W$ satisfying $\pi_\xi(\Gamma_W(x)) = 1$.

■

Before stating the next result, note that the open set $O_{x\,\xi}$ in Assumption 7 necessarily intersects C whenever $x \in C$. Hence, without loss of generality, we may assume that $O_{x\,\xi} \subset C$ for $x \in C$. We will use this fact in the proof of the following proposition:

Proposition 2.

Every invariant set in $M \times W$ under the Markov chain $\{(x_n, \xi_n)\}$ has positive $m_C \times \pi_\xi$ measure (Theorem 2.8.1, Condition A).

In fact, there exists an open set $V \subset C$ such that, for any invariant set I, the set $\Gamma \subset I$ (Γ depends on I) defined in Lemma 3 satisfies

$$m_C \times \pi_\xi (\Gamma \cap (V \times Q)) = m_C \times \pi_\xi (V \times Q) > 0.$$

In particular, any invariant set has full $m_C \times \pi_\xi$ measure in $V \times Q$.

Proof.

Let $I \subset M \times W$ be invariant and $\Gamma \subset I$ be as in Lemma 3.

Pick some $(x', \xi') \in C \times \text{int } Q$. Then, using Assumptions 6 and 7, Theorem 4.1.1 (b) (with the Decoupling Argument) shows that there must exist $n > 0$ and open sets $U \subset C$, $V \subset C$, and $R \subset Q$, with $x' \in U$ such that

$$P((x_n, \xi_n) \in B \times \text{int } Q \mid (x_0, \xi_0) = (x, \xi)) > 0$$

for all $B \in \mathcal{B}(M)$ such that $m_M (B \cap V) = m_C (B \cap V) > 0$ and for all $(x, \xi) \in U \times R$. Pick a point $(y, \zeta) \in I \cap (U \times R)$ (this set is not empty by Lemma 3). Then it must be that $m_C (\Gamma_M \cap V) = m_C(V)$ since, if not,

$$P((x_n, \xi_n) \in ((\Gamma_M)^c \cap V) \times Q \mid (x_0, \xi_0) = (y, \zeta)) > 0$$

and I is not invariant. Then, the set $\Gamma \equiv \{(x, \xi) \in M \times W : x \in \Gamma_M, \xi \in \Gamma_W (x)\} \subset I$ satisfies

$$m_C \times \pi_\xi (\Gamma \cap (V \times Q)) = \int_{\Gamma_M \cap V} \pi_\xi (\Gamma_W (x) \cap Q)\, m_C (dx)$$

$$= m_C (\Gamma_M \cap V) = m_C (V) = m_C \times \pi_\xi (V \times Q).$$

■

All of our subsequent results in this subsection will be built around the additional assumption that the maximal invariant control set $C \subset M$ is compact. Some statements can even be further strengthened under the assumption that M is compact (as in Proposition 1). In either case, the measure m_C is then finite and henceforth will be assumed to be standardized to a probability measure on C.

Proposition 3.

Under Assumption 4 (existence of a unique invariant probability measure for the $\{\xi_n\}$ process) and if C is compact, the pair Markov process $\{(x_n, \xi_n)\}$ has an invariant probability measure on $M \times W$.

Proof.

Pick a collection $\{O_i : i \geq 1\}$ of open and precompact sets such that $Q \subset \bigcup_{i=1}^{\infty} O_i$.

Let $A_j \equiv O_j \setminus \bigcup_{k=1}^{j-1} O_k$ so that $Q \subset \bigcup_{j=1}^{\infty} A_j$ with $A_i \cap A_j = \phi$ for $i \neq j$. We may also assume that, for all j, $\pi_\xi(A_j) > 0$ (otherwise this A_j will play no role in the ensuing argument and we may drop it). Since $\{\xi_n\}$ is ergodic, Theorem 2.8.2 implies that, π_ξ-a.e. and for $B \in \mathcal{B}(W)$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mu(k, \xi, B) = \pi_\xi(B).$$

Pick ξ outside the exceptional set over which the above limit may fail and define, for $x \in C$,

$$\lambda_{j\,n}(.) \equiv (n\,c_{j\,n})^{-1} \sum_{k=1}^{n} \bar{\mu}(k, (x, \xi), . \cap (C \times A_j)),$$

where $\bar{\mu}(k, (x, \xi), B)$ represents the k–step transition probability from (x, ξ) to $B \in \mathcal{B}(M \times W)$ and (using the Decoupling Argument)

$$c_{j\,n} \equiv \frac{1}{n} \sum_{k=1}^{n} \bar{\mu}(k, (x, \xi), C \times A_j) = \frac{1}{n} \sum_{k=1}^{n} \mu(k, \xi, A_j).$$

For each (n, j), $\lambda_{n\,j}$ is a probability measure (on $C \times A_j$) and, since, for all $n \in \mathbb{N}$ and any given j, $\text{supp}\,(\lambda_{n\,j}) = C \times \overline{A_j}$ is compact, the sequence $\{\lambda_{n\,j}\}$, j fixed, is tight. Therefore, by Helley–Bray's Theorem (see, e.g., Chung (1974, Theorem 4.3.3)), a subsequence of $\{\lambda_{n\,j}\}$ converges to a probability measure λ_j. Still using $n \in \mathbb{N}$ to represent this subsequence, we have:

$$\lambda_j(.) = (\lim_{n\to\infty} c_{j\,n})^{-1} \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \bar{\mu}(k, (x, \xi), . \cap (C \times A_j))$$

$$= (\pi_\xi(A_j))^{-1} \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \bar{\mu}(k, (x, \xi), . \cap (C \times A_j)).$$

Define $\lambda(.) \equiv \sum_{j=1}^{\infty} \pi_\xi(A_j)\,\lambda_j(.)$. Then

$$\lambda(C \times Q) = \sum_{j=1}^{\infty} \pi_\xi(A_j)\,(\pi_\xi(A_j))^{-1} \left[\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \bar{\mu}(k, (x, \xi), C \times A_j)\right]$$

$$= \sum_{j=1}^{\infty} \left[\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mu(k, \xi, A_j)\right]$$

$$= \sum_{j=1}^{\infty} \pi_{\xi}(A_j) = 1$$

So, λ is a probability measure on $C \times Q$ which can be extended to $M \times W$ simply by setting $\lambda((C \times Q)^c) = 0$. To complete the proof, it suffices to show that λ is invariant.

Let $z \equiv (x, \xi) \in M \times W$. The invariance of λ for the pair process $\{(x_n, \xi_n)\}$ follows from, for $A \in \mathcal{B}(M \times W)$ with $B = A \cap (C \times Q)$,

$$\int_{M \times W} \bar{\mu}(z, B)\, \lambda(dz) = \int_{C \times Q} \bar{\mu}(z, B)\, \lambda(dz)$$

$$= \int_{C \times Q} \bar{\mu}(z, B) \sum_{j=1}^{\infty} \pi_{\xi}(A_j)\, \lambda_j(dz)$$

$$= \int_{C \times Q} \bar{\mu}(z, B) \sum_{j=1}^{\infty} \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} 1_{[C \times A_j]}(z)\, \bar{\mu}(k, z', dz)$$

$$= \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \int_{C \times Q} \bar{\mu}(z, B)\, \bar{\mu}(k, z', dz)$$

$$= \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \bar{\mu}(k+1, z', B)$$

$$= \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \bar{\mu}(k, z', B)$$

$$= \sum_{i=1}^{\infty} \frac{\pi_{\xi}(A_i)}{\pi_{\xi}(A_i)} \left[\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \bar{\mu}(k, (x, \xi), B \cap (C \times A_i)) \right]$$

$$= \sum_{i=1}^{\infty} \pi_{\xi}(A_i)\, \lambda_i(B)$$

$$= \lambda(B)$$

Since the above reasoning also shows that $\int_{M \times W} \bar{\mu}(z, A \cap (C \times Q)^c)\, \lambda(dz) = 0$, we conclude that $\int_{M \times W} \bar{\mu}(z, A)\, \lambda(dz) = \lambda(B) = \lambda(A)$.

This completes the proof. ■

Remark 3.

One should note that the proof of Proposition 3 could have been approached via a result of Mañé (1987, Proposition 8.1) stating that if X is a compact space and $T : X \longrightarrow X$ is continuous, then the set of invariant probability measure for T is not empty. A proof of Proposition 3 based on this result can be found in Kliemann (1979) (since $Q = \text{supp}(\pi_{\xi})$ need not be compact, one has to use the fact that the process $\{\xi_n\}$ is ergodic). Also note that the proof of Proposition 3 only requires that the $\{\xi_n\}$ process possesses an invariant probability measure. Hence, under the compactness of C, the uniqueness of the invariant probability measure for the Markov chain $\{(x_n, \xi_n)\}$ is (together with its consequences) the only result requiring Assumptions 5 through 8. Both existence and finiteness of an invariant measure for this process are immediate.

Corollary 1.

If C is compact, every invariant set in M × W under $\{(x_n, \xi_n)\}$ is properly essential (Theorem 2.8.1, Condition B).

Proof.

By Proposition 3 and Remark 2.8.1 (3) (or Jain and Jamison (1967, Corollary 4.3)), the state space of the $\{(x_n, \xi_n)\}$ process, M × W, must be properly essential. Since, a direct application of Proposition 1, shows that M × W cannot contain two disjoint invariant sets (i.e., M × W is indecomposable), it necessarily follows that all invariant sets are properly essential (Chung (1964, Proposition 18.1)). ■

Armed with Conditions (A) and (B), we can now restate Theorem 2.8.1 in our context.

Theorem 1.

Let $\{(x_n, \xi_n)\}$ be the Markov process described at the beginning of Subsection 4.1. Then under Assumptions 1 through 8 and the compactness of C, the following statements are true:

a) There exists a unique (up to sets of $m_C \times \pi_\xi$ measure zero) $(m_C \times \pi_\xi)$-minimal invariant set $I' \subset M \times W$ for the Markov chain $\{(x_n, \xi_n)\}$.

b) When restricted to some state space $I \subset I'$ with $m_C \times \pi_\xi(I) = m_C \times \pi_\xi(I')$, the pair process $\{(x_n, \xi_n)\}$ becomes $(m_C \times \pi_\xi)$-recurrent, i.e., for all $B \in \mathcal{B}(M \times W)$ and for all $(x, \xi) \in I$,

$$m_C \times \pi_\xi(B \cap I) > 0 \text{ implies } L((x, \xi), B) = Q((x, \xi), B) = 1.$$

c) With the invariant set I is associated a unique (up to a multiplicative constant) invariant probability measure π such that, restricted to I,

$$m_C \times \pi_\xi << \pi.$$

Hence $\{(x_n, \xi_n)\}$ is ergodic.

Proof.

a) Under Conditions (A) and (B), Theorem 2.8.1 (a) guarantees the existence of a countable family of disjoint $(m_C \times \pi_\xi)$-minimal invariant sets $\{I_k ; k \in \mathbb{N}\}$. Then, by Proposition 2, eack I_k must have full measure in some set $V \times Q$, $V \subset C$, V open. It immediately follows that this family of disjoint invariant sets $\{I_k ; k \in \mathbb{N}\}$ must reduce (up to sets of $m_C \times \pi_\xi$ measure zero) to a single set, say I'.

b) This is an immediate application of Theorem 2.8.1 (b).

c) Uniqueness of a σ–finite invariant measure results from an immediate application of Theorem 2.8.1 (c). That this unique invariant measure is finite (i.e., can be standardized to a probability measure) follows from Proposition 3 since the finite invariant measure constructed there must be the unique invariant measure for the pair process $\{(x_n, \xi_n)\}$. The existence of a unique invariant probability measure then means (see Remark 2.8.3 (2)) that the process $\{(x_n, \xi_n)\}$ is ergodic. ■

Remark 4.

When $\{\xi_n\}$ is a sequence of iid random variables, Meyn and Caines (1988) use Orey (1971, Theorem 8.2 and the Corollary to Theorem 9.1) to obtain a result very similar to Theorem 1 above.

Theorem 2.

Write $(m_C \times \pi_\xi)\big|_I$ for the restriction of the product measure $m_C \times \pi_\xi$ to the Harris set I of Theorem 1. Then

a) The Markov chain $\{(x_n, \xi_n)\}$ on $C \times W$ is $(m_C \times \pi_\xi)\big|_I$-irreducible, weakly recurrent (with respect to this irreducibility measure) and its invariant probability measure π is equivalent to the maximal irreducibility measure for this process.

b) The Markov chain $\{(x_n, \xi_n)\}$ on $C \times W$ is π-irreducible and weakly recurrent (with respect to π). Moreover, supp $(\pi) = C \times Q$, and the unique invariant set is actually $C \times Q$ (up to sets of π measure zero).

Proof.

a) Part (b) of Theorem 1 states that, when restricted to I, we have a Harris recurrent chain (with respect to the reference measure $m_C \times \pi_\xi$), i.e., for all $B \in \mathcal{B}(C \times W)$ and for all $(x, \xi) \in I$,

$$m_C \times \pi_\xi (B \cap I) > 0 \Rightarrow L((x, \xi), B) = Q((x, \xi), B) = 1.$$

This automatically implies that, restricted to I, the Markov chain $\{(x_n, \xi_n)\}$ is $(m_C \times \pi_\xi)$-irreducible, i.e., $L((x, \xi), B) > 0$ for all $B \in \mathcal{B}(C \times W)$ and for all $(x, \xi) \in I$ such that $m_C \times \pi_\xi (B \cap I) > 0$.

Next, we show this holds for all $(x, \xi) \in C \times W$.

Fix $(x', \xi') \in I$. Then, using Assumptions 6 and 7, Theorem 4.1.1 shows that there must exist $n > 0$ and open sets $U \subset C$, $V \in C$, and $R \subset Q$, with $x' \in U$ such that

$$P((x_n, \xi_n) \in A \times Q \mid (x_0, \xi_0) = (x, \xi)) > 0$$

for all $A \in \mathcal{B}(M)$ such that $m_M(A \cap V) > 0$ and all $(x, \xi) \in U \times R$, i.e., for all $(x, \xi) \in U \times R$, $L((x, \xi), A \times Q) > 0$. Since $m_C \times \pi_\xi(I) = m_C \times \pi_\xi(I')$, I must have full measure in $V \times Q$ (see the argument in the proof of Proposition 2) and therefore $L((x, \xi), I) > L((x, \xi), I \cap (V \times Q)) > 0$ for all $(x, \xi) \in U \times R$.

Let $(y, \zeta) \in C \times W$ be arbitrary. By Proposition 1, $L((y, \zeta), U \times R) > 0$.

Then, by Lemma 1, we conclude that $L((y, \zeta), I) > 0$ for all $(y, \zeta) \in C \times W$

That we also have weak recurrence with respect to the reference measure $(m_C \times \pi_\xi)\big|_I$ is an immediate consequence of Theorem 2.7.2.

Then Theorem 2.8.2 (a) implies that π must be equivalent to the maximal irreducibility measure for the pair process $\{(x_n, \xi_n)\}$.

b) First note that π-irreducibility follows trivially from the fact that π is equivalent to the maximal irreducibility measure (see part (a)). Weak recurrence then follows from Theorem 2.7.2.

To show that supp $(\pi) = C \times Q$, we will show that, for all open sets $O \subset C \times Q$, $\pi(O) > 0$.

Suppose that, for some $O \subset C \times Q$, $\pi(O) = 0$. Then it follows that, for the Markov chain $\{(x_n, \xi_n)\}$ restricted to the invariant set I, on which $\{(x_n, \xi_n)\}$ is strongly recurrent ($(m_C \times \pi_\xi)$-recurrent), $O \cap I$ is inessential (see Lemma 2.3 in Jain and Jamison (1967)). Then, since $I = K \cup S$, $K \cap S = \phi$, where K is invariant and S is inessential or improperly essential with $\pi(S) = 0$ (see Theorem 2.7.2), we may also write

$$I = (K \setminus (O \cap I)) \cup (S \cup (O \cap I)).$$

But, since $K \subset I = \phi$, this would mean that $K \setminus O$ is invariant, which is impossible by Proposition 1. Hence, $\pi(O) > 0$ for all open sets O in $C \times Q$ and therefore supp $(\pi) = C \times Q$.

Finally, since the process $\{(x_n, \xi_n)\}$ is π-irreducible, one necessarily has that $\pi(I) = \pi(I') = \pi(C \times Q) < \infty$ (otherwise, I' cannot be invariant). ■

Corollary 2.

If M is compact, Theorem 2 holds for the Markov chain $\{(x_n, \xi_n)\}$ on $M \times W$.

Proof.

Simply repeat the proof of Theorem 2, using M instead of C. ■

Remark 5.

Since, by Theorem 2 (b), the pair process $\{(x_n, \xi_n)\}$ is π-irreducible and weakly recurrent on $C \times W$ (or, if M compact, on $M \times W$, by Corollary 2), Proposition 3.6 in Tweedie (1976) (see Remark 2.8.2) states that there exists a π-null set K such that, on the reduced state space $(M \times W) \setminus K$, the pair process is π-recurrent (i.e., strongly recurrent). (Note that, under the compactness of C, this holds even if M is not compact since $(C \times W)^C$ is a π-null set.) Hence, for all $(x, \xi) \in (M \times W) \setminus K$, $L((x, \xi), B) = 1$ for all $B \in \mathcal{B}(M \times W)$ such that $\pi(B) > 0$.

Since, by the proof of Theorem 2 (b), all open sets in $C \times Q$ have positive π measure, this implies that $P([S^n_{x\,\xi}, \xi_n] \in O \text{ i.o.} \mid (x_0, \xi_0) = (x, \xi)) = 1$ π-a.e. for any open set $O \subset C \times Q$, where $S^n_{x\,\xi}$ denotes the position at the n^{th} step of the trajectory starting from (x, ξ). We can also write, for any open set $U \subset C$,

$$P(S^n_{x\,\xi} \in U \text{ i.o.} \mid (x_0, \xi_0) = (x, \xi))$$

$$\equiv P\left([S^n_{x\,\xi}, \xi_n] \in U \times W \text{ i.o.} \mid (x_0, \xi_0) = (x, \xi)\right) = 1 \ \pi\text{–a.e.},$$

which means that, if C is compact, the orbit of any point $x \in M$ satisfying $(x, \xi) \in (M \times W) \setminus K$ will enter any open set in C infinitely often with probability one.

So far, in the above remark, we have established that, if C (or M) is compact, with probability one and in the M component, the orbit of the pair process enters any open subset of the maximal invariant control set, except maybe for initial values (x, ξ) in a π-null set. Still, this is not extremely useful since, in this setup, $(M \times Q) \setminus C \times Q$ (recall that $C \times Q$ is the support of π) has π measure zero and could be quite large. Fortunately, more can be said.

Theorem 3.

Let M be compact and assume that the $\{\xi_n\}$ process is π_ξ-recurrent (strongly recurrent). Then, the pair process $\{(x_n, \xi_n)\}$ is π-recurrent (strongly recurrent) on $M \times W$.

In particular, for any set $B \subset C$ of positive m_C measure, we have that $P(S^n_{x\,\xi} \in B \text{ i.o.} \mid x_0 = x, \xi_0 = \xi) = 1$ for all initial values $(x, \xi) \in M \times W$.

Proof.

Mimicking an argument taken from Meyn (1989, Proposition 3.1), we will first show that, for all $(x, \xi) \in M \times W$, for all open sets $O \subset C$, and for all compact sets $K \subset M$,

$$P\left[\left[S_{x\,\xi} \cap O \neq \phi \mid (x_0, \xi_0) = (x, \xi)\right] \cup \left[S^n_{x\,\xi} \in K \text{ i.o.} \mid (x_0, \xi_0) = (x, \xi)\right]^c\right] = 1.$$

Since the compact set K is arbitrary, taking K = M will then give that

$$P\left[\left[S^n_{x\,\xi} \in M \text{ i.o.} \mid (x_0, \xi_0) = (x, \xi)\right]^c\right] = 0,$$

and we will be able to conclude that, for all $(x, \xi) \in M \times W$,

$$P\left[\left[S^n_{x\,\xi} \in O \text{ i.o.} \mid (x_0, \xi_0) = (x, \xi)\right]\right] = 1.$$

Let $O \subset C$ be open and recall that, by Proposition 1,

$$L((x, \xi), O \times W) > L((x, \xi), O \times \text{int } Q) > 0,$$

i.e., for some $k \geq 1$ and some $\delta_{x\,\xi} > 0$, $\bar{\mu}(k, (x, \xi), O \times \text{int } Q) = 2\,\delta_{x\,\xi} > 0$. Since $\{(x_n, \xi_n)\}$ is Feller (Proposition 4.1.1) and $O \times \text{int } Q$ is an open set in $M \times W$, the function $\bar{\mu}(k, ., O \times \text{int } Q)$ is lower semi–continuous (Lemma 4.1 in Cogburn (1975)) where $\bar{\mu}(.,.)$ is the kernel of the pair process $\{(x_n, \xi_n)\}$.

This implies that there exists an open neighborhood of x, say U_x, such that $\bar{\mu}(k, (y, \xi), O \times \text{int } Q) > \delta_{x\,\xi}$ for all $y \in U_x$, i.e., we have $L((y, \xi), O \times \text{int } Q) > \delta_{x\,\xi}$ for all $y \in U_x$.

Pick an arbitrary compact set $K \subset M$. Since K can be covered by a finite collection of open sets, it follows that there exists $\delta_K > 0$ such that, for every $x \in K$,
$L((x, \xi), O \times W) > L((x, \xi), O \times \text{int } Q) > \delta_K$.

Now, by Proposition 5.1 in Orey (1971), $L((x, \xi), O \times W) > \delta_K$ for all $x \in K$ implies that, P-a.s.,

$$\Lambda(K) \equiv \left[\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \left[(x_k, \xi_k) \in K \times W\right] \;\middle|\; (x_0, \xi_0) = (x, \xi)\right]$$

$$\subset \Lambda(O) \equiv \left[\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \left[(x_k, \xi_k) \in O \times W\right] \;\middle|\; (x_0, \xi_0) = (x, \xi)\right],$$

for all $(x, \xi) \in M \times W$, i.e., that

$$\left[S^n_{x\,\xi} \in K \text{ i.o. } | \ (x_0, \xi_0) = (x, \xi)\right] \subset \left[S^n_{x\,\xi} \in O \text{ i.o. } | \ (x_0, \xi_0) = (x, \xi)\right],$$

for all $(x, \xi) \in M \times W$.

So, $P\left[\Lambda(O) \cup (\Lambda(K))^c\right] = 1$ for all $(x, \xi) \in M \times W$, and hence,

$P\left[\left[S^n_{x\,\xi} \in O \text{ i.o. } | \ (x_0, \xi_0) = (x, \xi)\right] \cup \left[S^n_{x\,\xi} \in K \text{ i.o. } | \ (x_0, \xi_0) = (x, \xi)\right]^c\right] = 1$ for all $(x, \xi) \in M \times W$. This completes the argument discussed at the onset of this proof.

So far, we have shown that, from any $(x, \xi) \in M \times W$, the pair process $\{(x_n, \xi_n)\}$ enters $O \times W$ (and hence, using the strong recurrence of the $\{\xi_n\}$ process, $O \times T$, for any open set $T \subset \text{int } Q$) with probability one. But this is not sufficient to ensure π-recurrence. Indeed, the Harris set $I \subset C \times Q$ only satisfies $\pi(I) = \pi(C \times Q)$ and entering $O \times T$ for any open sets $O \subset C$ and $T \subset \text{int } Q$ does not guarantee that (at least in the M component) one enters I with probability one. That the M component of the pair process $\{(x_n, \xi_n)\}$ does enter I with probability one is what we prove next. Pick $x \in C$ and let $U \subset C$, $V \subset C$, and $R \subset \text{int } Q$ be open sets such that $x \in U$, $\xi \in R$, and for, some $c > 0$,

$$P(x_n \in B \cap V) | \ (x_0, \xi_0) = (y, \zeta)) \geq c\, m_C(B \cap V),$$

for all $B \in \mathcal{B}(M)$, all $(y, \zeta) \in U \times R$. The existence of such open sets U, V, and R is guaranteed by Theorem 4.1.1 (b).

By Theorem 2.7.2, I is invariant and hence, by Lemma 3, I contains a set of the form $\Gamma = \{(x, \xi) \in M \times W : x \in \Gamma_M, \xi \in \Gamma_W(x)\}$ ($\Gamma_M \subset C$, $\Gamma_W(x) \subset W$ with, for all $x \in \Gamma_M$, $\pi_\xi(\Gamma_W(x)) = 1$) which has full measure in $V \times Q$.

This implies that $L((y, \zeta), \Gamma_M \times Q) > \alpha > 0$ for all $(y, \zeta) \in U \times R$ (α depends on ξ and x through U and V).

But, the Decoupling Argument, the invariance of int Q, and the fact that, for all $\xi \in$ int Q, $\mu(\xi, B) = 0$ for all $B \in \mathcal{B}(W)$ such that $\pi_\xi(B) = 0$ (see the discussion of Assumption 6) imply that

$$L((y, \zeta), I) \geq L((y, \zeta), \Gamma) = L((y, \zeta), \Gamma_M \times Q) > \alpha.$$

Indeed, for $x_n = x \in \Gamma_M$ and $\xi_{n-1} = \xi \in$ int Q, $P(\xi_n \in \Gamma_W(x) \mid \xi_{n-1} = \xi) = 1$ since $\pi_\xi(\Gamma_W(x)) = 1$ for all $x \in \Gamma_M$.

Therefore, again by Proposition 5.1 in Orey (1971), for all $(x, \xi) \in M \times W$,

$$P([S^n_{y\,\xi}, \xi_n] \in I \text{ i.o.} \mid (x_0, \xi_0) = (x, \xi))$$

$$\geq P([S^n_{x\,\xi}, \xi_n] \in U \times R \text{ i.o.} \mid (x_0, \xi_0) = (x, \xi)) = 1.$$

Since the process $\{(x_n, \xi_n)\}$ is π-recurrent on I, this shows that it is π-recurrent on $M \times W$. In particular, if $m_C(B) > 0$, then $m_C \times \pi_\xi(B \times Q) > 0$, $\pi(B \times Q) > 0$ and hence we have that, for all initial values $(x, \xi) \in M \times W$,

$$P(S^n_{x\,\xi} \in B \text{ i.o.} \mid x_0 = x, \xi_0 = \xi) = L((x, \xi), B \times W) = 1,$$

and this completes the proof. ■

<u>Remark 6.</u>

1) The proof of the Theorem 3 shows that strong recurrence of the $\{\xi_n\}$ process as well as compactness of C (without compactness of M) are sufficient to prove π-recurrence on $C \times W$ (simply replace M by C in the proof of Theorem 3.).

2) The last statement of Theorem 3 means that we have, loosely speaking, "strong recurrence" in the M component, i.e., that, if $B \subset M$ is such that $\pi(B \times Q) > 0$, then $L((x, \xi), B \times W) = 1$. Also note that, from the proof of Theorem 3, this holds true even if $\{\xi_n\}$ itself is not strongly recurrent. In other words, in order to obtain a strongly recurrent pair process from the weakly recurrent pair process $\{(x_n, \xi_n\})$, is suffices to reduce the state space of the $\{\xi_n\}$ process to $W \setminus A$, for some appropriate set $A \subset W$ with $\pi_\xi(A) = 0$ (see Remark 2.8.2).

Before proving our last result in this section, it is useful to make a short remark about cycles.

Remark 7.

By Theorem 2.8.3 (b) (or, for example, Theorem 2.2 in Jain and Jamison (1967)), the unique (again up to a π-null set) invariant set $C \times Q$ can possibly be decomposed into a cycle $\{C_i ; 1 \leq i \leq d\}$, for some $d \geq 1$. Under Assumption 5, the $\{\xi_n\}$ process itself is necessarily aperiodic and we may therefore write $C_i = C_i' \times Q$ and

$$C \times Q = \bigcup_{i=1}^{d} (C_i' \times Q),$$

where the C_i' sets are disjoint. Moreover, we must have $C_i' = \overline{C_i'}$ for $1 \leq i \leq d$ since, if C_{i_0}' is not closed for some i_0, continuity of $f(., \xi)$ would imply that none of these C_i' sets is closed and neither should be their disjoint union, which does contradict the closedness of C. From these facts, it follows that if we have such a cycle, it must be that C is not connected. (Also see Proposition 3.1.11 and Corollary 2.2.b in Meyn (1989).)

Theorem 4.

a) If M is compact and $\{\xi_n\}$ is strongly recurrent, we have, for any initial distribution ν on $\mathcal{B}(M \times W)$ and with $\|.\|$ denoting the total variation,

$$\lim_{n\to\infty} \left\| \frac{1}{n} \int_{M \times W} \nu(dy \times d\xi) \sum_{k=1}^{n} \bar{\mu}(k, (y, \xi), .) - \pi(.) \right\| = 0,$$

where $\bar{\mu}(.,.)$ denotes the kernel of the pair process $\{(x_n, \xi_n)\}$ and π is the unique invariant probability measure for $\{(x_n, \xi_n)\}$.

In particular, for all $B \in \mathcal{B}(M \times W)$ and independently of $(x, \xi) \in M \times W$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \bar{\mu}(k, (x, \xi), B) = \pi(B),$$

b) Moreover, if the pair process is aperiodic, then,

$$\lim_{n\to\infty} \left\| \int_{M \times W} \bar{\mu}(n, (y, \xi), .)\, \nu(dy \times d\xi) - \pi(.) \right\| = 0,$$

which implies that, for all $B \in \mathcal{B}(M \times W)$ and independently of $(x, \xi) \in M \times W$,

$$\lim_{n\to\infty} \bar{\mu}(n, (x, \xi), B) = \pi(B).$$

c) If the unique invariant set can be decomposed into a cycle $\{C_n' \times Q\ ;\ 1 \le n \le d\}$, $d > 1$, then there is a unique collection of probability measures $\{\pi_n\ ;\ 1 \le n \le d\}$ with $\pi_i(C_j' \times Q) = 0$ for $i \neq j$, $\pi_i(C_i' \times Q) = 1$ and such that for all sets $B \in \mathcal{B}(M \times W)$,

$$\lim_{n\to\infty} \bar{\mu}(nd+m, (x, \xi), B) = \pi_k(B), \quad (x, \xi) \in C_i' \times Q \text{ and } k = i + m \pmod d.$$

Moreover, the unique invariant probability measure π satisfies, for all $(x, \xi) \in M \times W$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \bar{\mu}(k, (x, \xi), B) = d^{-1} \sum_{n=1}^{d} \pi_n(B),$$

and, for all initial distribution ν on $\mathcal{B}(M \times W)$, we have

$$\lim_{n\to\infty} \left\| \frac{1}{d} \int_{M \times W} \sum_{k=1}^{d} \bar{\mu}(n+k, (y, \zeta), .)\, \nu(dy \times d\zeta) - \pi(.) \right\| = 0.$$

Proof.

a) By Theorem 3, the pair process $\{(x_n, \xi_n)\}$ is π-recurrent (strongly recurrent) on $M \times W$. Since, by Theorem 1, $\{(x_n, \xi_n)\}$ is ergodic, the limit statement given is an immediate application of Theorem 2.8.2 (d) and Remark 2.8.3 (2).

b) Since we have a π–recurrent and ergodic process (see part (a)), this is an immediate consequence of Theorem 2.8.3 (a).

c) Since we have a π–recurrent and ergodic process (see part (a)), this is an immediate consequence of Theorem 2.8.3 (b). ■

Remark 8.

1) If $\{\xi_n\}$ is only weakly recurrent, the results of Theorem 4 hold for all $(x, \xi) \in M \times (W \setminus A)$ where A is a π_ξ-null set (see Remark 6 (2)).

2) If M is not compact but C is compact, the results of Theorem 4 hold for all $(x, \xi) \in C \times W$) (when $\{\xi_n\}$ is strongly recurrent) or, if $\{\xi_n\}$ is only weakly

recurrent, for all $(x, \xi) \in C \times (W \setminus A)$, with $\pi_{\xi}(A) = 0$ (see Remarks 6 (1) and 6 (2)).

In this section and under the assumptions specified in Subsection 4.1 and the compactness of C, we have therefore established the existence of a unique invariant probability measure π for the Markov chain $\{(x_n, \xi_n)\}$, i.e., the ergodicity (positivity and weak recurrence) of this pair process. Moreover we have also shown that, if M is compact and $\{\xi_n\}$ is strongly recurrent, we have in fact strong recurrence (on $M \times W$) with respect to the reference measure π.

The next step will be to use these results in conjunction with Oseledeč's Multiplicative Ergodic Theorem to study the Lyapunov spectrum associated with the linear dynamical system defined by $x_{n+1} = A(\xi_n) x_n$. This is the topic of the next section.

5. STABILITY PROPERTIES OF $X_{n+1} = A(\xi_n) X_n$

5.1. The Lyapunov Spectrum

As the title of this section indicates, we will now be concerned with the stability properties of the linear system $x_{n+1} = A(\xi_n) x_o$ in $\mathbb{R}_0^d$. We will assume that:

a) $\{\xi_n\}$ is a Feller stationary ergodic process valued in some (connected) manifold W^k and with invariant (probability) measure π_ξ (and transition kernel $\mu(.,.)$),

b) $A(.)$ is a C^1 map from the manifold W into the Lie group $Gl(d, \mathbb{R})$ satisfying the condition $\sup \{\|A(\xi)\| ; \xi \in \text{supp}(\pi_\xi)\} = K < \infty$ (the norm is the operator norm), and

c) x_o is an $\mathbb{R}_0^d$ valued random variable independent of $\{\xi_n ; n \in \mathbb{N}\}$.

Remark 1.

The boundedness assumption on $A(.)$ will ensure that the moment Lyapunov exponents (defined in Subsection 5.3) as well as the almost sure Lyapunov exponents (see below) are finite (see the integrability condition in Oseledeč's Theorem (Theorem 2.9.1)). Moreover, in Subsection 5.1, $A(.)$ being a measurable map is sufficient, but further developments, starting in Subsection 5.2, will require $A(.)$ to be C^1 (compare with Assumption 2 in Subsection 4.1).

The starting point for our stability study will be the use of Oseledeč's Multiplicative Ergodic Theorem (Theorem 2.9.1). Hence we first describe how the

present setup fits into the framework of this result.

Let $\{\zeta_n ; n \in \mathbb{N}\}$ be an arbitrary stationary process (defined on some probability space $(\Omega', \mathcal{F}, P')$) taking values in W. It is well known (see Breiman (1968) and Doob (1953)) that, on its path space $\Omega \equiv W^{\mathbb{N}}$ and in terms of distributions (see below), any stationary process can be generated by a measure preserving transformation $\otimes$ (i.e., a measurable map from Ω to Ω satisfying, for all $B \in \mathcal{B}(\Omega) = \mathcal{B}(W^{\mathbb{N}})$, $P(\otimes^{-1}(B)) = P(B)$, with P denoting the measure on the path space Ω induced by the finite dimensional distributions of the $\{\zeta_n\}$ process. To do this, one uses the shift operators $\otimes$ on Ω, i.e., $\otimes$ is defined by

$$\otimes \omega = \otimes (\zeta_0, \zeta_1, \ldots, \zeta_n, \ldots) = (\zeta_1, \zeta_2, \ldots, \zeta_n, \ldots),$$

which we abbreviate by $\otimes \omega(.) = \otimes \omega(. + 1)$. The operator $\otimes$ then preserves the measure P. Writing H_n for $\otimes^n$ ($H_0 \omega \equiv \omega$) and defining the random variable $\zeta : \Omega \longrightarrow W$ on $(\Omega, \mathcal{B}(\Omega), P)$ by $\zeta(\omega) = \zeta_1$, the process defined by

$$\{\zeta(H_n \omega) ; n \geq 0\} \equiv \{\zeta_n(\omega) ; n \geq 0\}$$

is stationary and distributed as $\{\zeta_n\}$ (Breiman (1968, Proposition 6.9)). Moreover, if the transformation $\otimes$ is ergodic, i.e., if, for all $B \in \mathcal{B}(\Omega)$ satisfying the $\otimes$-invariance condition $B = \otimes^{-1}(B)$, $P(B) = 0$ or 1, then the process $\{\zeta_n(\omega)\}$ is also ergodic (Breiman (1968, p. 116)).

Remark 1.

In Subsection 4.1, $(\Omega, \mathcal{F}, P)$ was defined to be the underlying probability space for the Markov process $\{\xi_n ; n \in \mathbb{N}\}$ and in Subsection 2.9, path spaces were denoted by X. Nevertheless, a switch in notation is convenient because later expressions will

explicitely depend on the paths (and the shift operators H_n on these paths) while the use of Ω rather than X is meant to emphasize the fact that the randomness associated with our system arises from the probability measure P on the measurable space $(\Omega, \mathcal{B}(\Omega)) = (W^{\mathbb{N}}, \mathcal{B}(W^{\mathbb{N}}))$.

Also, when an expression or result depends crucially on the path ω (i.e., on the flow associated with the stationary process), this ω path dependency will be explicitely indicated. When only the finite dimensional distributions of the stationary process are relevant, we will use $\{\zeta_n\}$ (defined on $(\Omega', \mathcal{F}, P')$) or $\{\zeta_n(\omega)\}$ (defined on $(\Omega, \mathcal{B}(\Omega), P)$) interchangeably, since both processes then have the same finite dimensional distributions.

In terms of the discussion in Subsection 2.9 (see in particular the example of products of random matrices in Remark 2.9.4), we can write the current setup as follows:

On $\Omega \equiv W^{\mathbb{N}}$, we have the flow $H : \mathbb{N} \times \Omega \longrightarrow \Omega$ defined by

$$H_n \omega(.) = \omega(. + 1).$$

To this flow, we associate the cocycle $C : \mathbb{N} \times \Omega \longrightarrow Gl(d, \mathbb{R})$ defined by

$$C(n, \omega) = A(H_{n-1}\omega)\, A(H_{n-2}\omega) \ldots A(H_0\omega)$$

or, equivalently,

$$C(n, \omega) = A(\xi_{n-1}(\omega))\, A(\xi_{n-2}(\omega)) \ldots A(\xi_0(\omega)),$$

representing the linear stochastic difference equation.

Remark 2.

In Sections 4 to 6, the stationary process used is in fact the Markov process $\{\xi_n ; n \in \mathbb{N}\}$. Nevertheless the Markov character of the $\{\xi_n\}$ process has no bearing on the arguments of this section. Indeed, these arguments depend on the flow $\{H_n ; n \in \mathbb{N}\}$ associated with the stationary process $\{\xi_n(\omega)\}$ (on its path space Ω) and flow properties are not necessarily related to the Markov character of the $\{\xi_n\}$ process (this is true, in particular, for Oseledeč's Multiplicative Ergodic Theorem (Theorem 2.9.1)). The fact that the $\{\xi_n(\omega)\}$ process is actually Markov will only become important when, in Subsection 5.2, we start using the unique invariant probability measure for the pair process $\{(x_n |x_n|^{-1}, \xi_n(\omega))\}$, whose existence was, under additional assumptions specified at the end of this subsection (see also Subsection 4.1), demonstrated in Section 4.

Definition 2.9.10 now yields the Lyapunov exponents

$$\lambda(\omega, x) = \overline{\lim_{n\to\infty}} \frac{1}{n} \log \|C(n, \omega)\, x\|$$

with ω belonging to the new probability space $\Omega = W^{\mathbb{N}}$.

The assumptions of Oseledeč's Multiplicative Ergodic Theorem are then satisfied under our setup because:

1) As already mentioned above, due to the stationarity and ergodicity of the process $\{\xi_n(\omega) : n \in \mathbb{N}\}$, the measure P on Ω is H_n-invariant for all $n \in \mathbb{N}$, and ergodic.
2) By Assumption (b), $E(A(\xi_n)) < \infty$ for all $n \in \mathbb{N}$, i.e., Oseledeč's integrability condition is trivially satisfied.

We therefore obtain:

1) There exists a set $\Omega_0 \subset \Omega$, $P(\Omega_0) = 1$, such that, if $\omega \in \Omega_0$, then ω and the cocycle $C(n, \omega)$ are regular. Hence $\lambda(\omega, x) = \lim_{n \to \infty} \frac{1}{n} \log \|C(n, \omega) x\|$.

2) Associated with the system $x_{n+1} = A(\xi_n(\omega)) x_0$ in $\mathbb{R}^d_0$, there is only a finite number $p \leq d$ of different Lyapunov exponents $\lambda_1 > \ldots > \lambda_p$, with respective multiplicities $d_1, \ldots, d_p$ ($d_1 + \ldots + d_p = d$). Thanks to ergodicity, the values p, $\lambda_1, \ldots, \lambda_p$, and $d_1, \ldots, d_p$ are all nonrandom (i.e., do not depend on the ω path).

More can be said (refer to Subsection 2.9) but these two statements are the most important ones for the developments hereafter.

Remark 3.

At this point, it is important to stress that the above is true provided that the flow $\{H_n ; n \in \mathbb{N}\}$ is stationary and ergodic with respect to the measure P, which is true if the $\{\xi_n(\omega)\}$ process is stationary and ergodic (with the invariant measure π_ξ being the initial distribution for the process $\{\xi_n(\omega)\}$) and both H_n and P are constructed as at the beginning of this section. Other choices of the operators H_n and P statistically yielding the same process $\{\xi_n(\omega)\}$ (i.e., giving the same finite dimensional distributions for the $\{\xi_n(\omega)\}$ process) may correspond to different flows and hence give different Lyapunov exponents.

Now, according to the discussion in Subsection 2.9, the stochastic system $x_{n+1} = A(\xi_n(\omega)) x_0$ will be (exponentially) stable for all $x_0 \in M$ if $\lambda_1 < 0$. But, in general, the maximal growth behavior of such an equation will only be seen for initial

values x_0 in some subset of $\mathbb{R}_0^d$ (the subset $L_1 \setminus L_{1+1}$). It is therefore useful to establish conditions under which the deterministic Lyapunov spectrum of the system, $\{(\lambda_i, d_i)\,;\, 1 \le i \le p\}$, reduces to $\{(\lambda_1, d_1)\}$ with probability one, i.e., under which $p = 1$ with probability one. A result of this nature is the main trust of this section and is given in Theorem 5.2.1.

Remark 4.

Note that, even if $p = 1$, the multiplicity of the unique (with probability one) Lyapunov exponent may not be d. Other Lyapunov exponents may be present but they "show up" with probability zero.

In order to establish such a result and following the ideas of, for example, Arnold et al. (1986a), we will project our linear system onto $\mathbb{P}^{d-1}$ by defining $s_n = |x_n|^{-1} x_n$. The stochastic difference equation can then be written as

$$s_{n+1} = |A(\xi_n(\omega))\, s_n|^{-1} A(\xi_n(\omega))\, s_n$$

$$= f(s_n, \xi_n(\omega)),$$

where $f(s, \xi) \equiv |A(\xi)\, s|^{-1} A(\xi)\, s$.

The rationale for working with the projected system $s_{n+1} = f(s_n, \xi_n(\omega))$ arises from the following facts:

1) The logarithmic norm of a solution of the unprojected system at time n and starting at the random initial value (x_0, ξ_0), $\log |x(n, (x_0, \xi_0), \omega)|$, can be expressed as a function of the pair process $\{(s_n, \xi_n) : n \in \mathbb{N}\}$, namely,

$$\log |x(n, (x_0, \xi_0), \omega)| = \sum_{i=0}^{n} h(s_i, \xi_i) + \log |x_0|,$$

where $h(a, b) \equiv \log |A(b)\, a|$ (see the proof of Theorem 5.2.1).

2) For the deterministic control system associated with the stochastic system $x_{n+1} = A(\xi_n(\omega))\, x_n$, projection onto $\mathbb{P}^{d-1}$ (which leads to the nonlinear deterministic control system $s_{n+1} = f(s_n, u_n)$ (see the previous page and the discussion below)), will ensure the uniqueness of the maximal invariant control set under the condition that, for all $s \in \mathbb{P}^{d-1}$, int $Ss \neq \phi$ (see Theorem 3.3.1).

3) $\mathbb{P}^{d-1}$ is compact. This will be used to establish the existence of an invariant probability measure for the pair process $\{(s_n, \xi_n) : n \in \mathbb{N}\}$ (see Theorem 4.2.1). This will also be convenient to establish the uniqueness of that invariant probability measure (via uniqueness of the maximal invariant control set for the associated dynamical control system as in part (2) above). Note nevertheless that other approaches can be used to establish uniqueness of the maximal invariant control set: In the continuous time case, Colonius and Kliemann (1989, Proposition 2.4) use a global (on a compact set) asymptotic stability condition for steady state solutions of the associated control system.

As we have just hinted to, the existence of an invariant probability measure π for the Markov process $\{(s_n, \xi_n) : n \in \mathbb{N}\}$ will play a crucial role in the following section. In Subsection 4.1, we have given a series of assumptions which guarantee the existence of such a π. Some of these assumptions are already duplicated in Assumptions (a) through (c) at the beginning of this section, namely, Assumptions 1 through 4. Therefore, we only need to complete our list of working hypotheses

accordingly. Hence, we will further assume that:

d) For all $\xi \in \text{int } Q$ ($Q \equiv \text{supp } (\pi_\xi)$) and all sets $B \in \mathcal{B}(W)$, $\pi_\xi (B) > 0$ implies that $\mu (\xi, B) > 0$. Moreover, $\pi_\xi (Q) = \pi_\xi (\text{int } Q)$.

e) In the decomposition of the one–step transition probability $\mu (\xi, .)$ into its absolutely continuous and singular components (with respect to the volume element m_W on W),

$$\mu (\xi, B) = \int_B g (\xi, y)\, m_W (dy) + \mu_s (\xi, B)$$

for $B \in \mathcal{B}(W)$ and $\xi \in W$, the density function $g (\xi, .)$ is, for all $\xi \in \text{int } Q$, strictly positive on int Q and the map g is jointly lower semi–continuous.

(Recall that the Markov chain $\{\xi_n\}$ is now assumed to be stationary and that Assumption (e) then implies the first statement of Assumption (d).)

e) The deterministic semigroup arising from the control system $s_{n+1} = f (s_n, u_n)$, $u_n \in \text{int } Q$, associated with the projection on $\mathbb{P}^{d-1}$ of the stochastic dynamical system $x_{n+1} = A (\xi_n)\, x_n$, satisfies the condition that int $Ss \neq \phi$ for all $s \in \mathbb{P}^{d-1}$.

g) for every $(s, \zeta) \in \mathbb{P}^{d-1} \times W$ there exists a time $n_{s\,\zeta}$ and an open set $O_{s\,\zeta}$ included in $\mathbb{P}^{d-1}$ (with the subscripts indicating a dependence on the points s and ζ) such that $m_P (O_{s\,\zeta} \cap .) << P \left[s_{n_{s\zeta}} \in . \mid s_0 = s, \xi_0 = \zeta \right]$, where m_P denotes the volume element on $\mathbb{P}^{d-1}$.

Remark 5.

At this point, recall that, under Assumption (f), the deterministic control system $s_{n+1} = f (s_n, u_n)$ admits one and only one maximal invariant control set $C \subset \mathbb{P}^{d-1}$ given by

$$C = \overline{C} = \bigcap \{\overline{Ss} ; s \in \mathbb{P}^{d-1}\} \neq \phi$$

(see Theorem 3.3.1). C is therefore closed (and Borel), with nonempty interior (since int $Ss \neq \phi$ for all $s \in \mathbb{P}^{d-1}$).

Therefore, under Assumptions (a) through (g) and using the notation of Section 4 (with $M = \mathbb{P}^{d-1}$), the pair process $\{(s_n, \xi_n)\}$ admits a unique invariant probability measure $\pi >> m_C \times \pi_\xi$, whose support is $C \times Q$ (see Theorems 4.2.1 and 4.2.2).

This fact will be used in the next section to claim that the Lyapunov spectrum of the stochastic system $x_{n+1} = A(\xi_n(\omega))\, x_n$ reduces, with probability one, to a single Lyapunov exponent $\lambda(\omega, x) = \lambda_1$, the top (largest) Lyapunov exponent in the spectrum.

5.2. Sample Stability

In this subsection, we investigate the basic properties of the top Lyapunov exponent for the unprojected system $x_{n+1} = A(\xi_n)\, x_n$ on $\mathbb{R}^d_0$ via the projected system $s_{n+1} = f(s_n, \xi_n)$ on $\mathbb{P}^{d-1}$, $s_n = x_n\, |x_n|^{-1}$ and $f(s, \xi) = A(\xi)\, s\, |A(\xi)\, s|^{-1}$ (see Subsection 5.1). Recall that, from now on, A is assumed to be at least a C^1 map from $W^k \longrightarrow Gl(d, \mathbb{R})$ and that we denote by π_ξ, π, and P the invariant initial distribution for the $\{\xi_n(\omega)\}$ process, the invariant probability measure for the pair process $\{(s_n, \xi_n(\omega))\}$, and the probability measure on the path space $\Omega \equiv W^{\mathbb{N}}$

generated by the transition probabilities of the $\{\xi_n\}$ process, respectively.

Theorem 1.

Under Assumptions (a) through (g) of Subsection 4.1, the Lyapunov's spectrum of the system $x_{n+1} = A(\xi_n(\omega))\, x_n$, $x_0 \neq 0$, contains, P-a.s., only one (nonrandom) value $\lambda(\omega, x) = \lambda_1$ which is independent of the value $x \in \mathbb{R}_0^d$ of the random variable x_0. Moreover, we have, P-a.s.,

$$\lambda_1 = E_\pi \left[\log \left| A(\xi_0(\omega)) \frac{x}{|x|} \right|\right] = \lim_{n\to\infty} \frac{1}{n} \log \|A(\xi_n(\omega)) \ldots A(\xi_0(\omega))\|.$$

Proof.

Write $x(n, (x, \xi_0), \omega) \equiv A(\xi_n(\omega)) \ldots A(\xi_0(\omega))\, x$ for the solution at time n of the stochastic equation $x_{n+1} = A(\xi_n(\omega))\, x_n$ with initial value $x \in \mathbb{R}_0^d$. We want to compute

$$\lim_{n\to\infty} \frac{1}{n} \log |x(n, (x, \xi_0), \omega)|.$$

Note that this limit exists P-almost surely since, by Oseledeč's Theorem, there exists $\Omega_0 \subset \Omega$ with $P(\xi_0(\omega) \in \Omega_0) = 1$ and such that $\omega \in \Omega_0$ implies that ω is regular. As in Subsection 5.1, we will write $s_n = |x_n|^{-1} x_n$ and $h(a, b) = \log |A(b)\, a|$. We then have (dropping the ω's for convenience):

$$\log |x(n, (x, \xi_0), \omega)| = \log |A(\xi_n) \ldots A(\xi_0)\, x|$$

$$= \log \frac{|A(\xi_n) \ldots\ldots A(\xi_0)\, x|}{|A(\xi_{n-1}) \ldots A(\xi_0)\, x|}$$

$$+ \log \frac{|A(\xi_{n-1}) \ldots A(\xi_0)\, x|}{|A(\xi_{n-2}) \ldots A(\xi_0)\, x|} + \ldots + \log \frac{|A(\xi_0)\, x|}{|x|}$$

$$+ \log |x|$$

$$= \sum_{i=1}^{n} \log \left| A(\xi_i) \frac{A(\xi_{i-1}) \dots A(\xi_0)\, x}{|A(\xi_{i-1}) \dots A(\xi_0)\, x|} \right|$$

$$+ \log |x|$$

$$= \sum_{i=1}^{n} h(s_i, \xi_i) + \log |x|$$

Since $h(.,.)$ is a measurable and π-integrable function of the ergodic process $\{(s_n, \xi_n)\}$, by the Birkhoff Ergodic Theorem (see Doob (1953, Theorem 7.6.2)), we have, P-a.s.,

$$\lim_{n\to\infty} \frac{1}{n} \log |x(n, (x, \xi_0), \omega)| = E_\pi(h(s, \xi_0(\omega)))$$

$$= E_\pi(\log |A(\xi_0(\omega))\ s|) \quad \pi\text{-a.e.} \qquad (*)$$

Therefore and even for a random initial value x_0, this P-almost sure limit does not depend on the initial value of the $\{(s_0, \xi_0(\omega))\}$ process (nor does it depend on the initial value of the $\{(x_0, \xi_0(\omega))\}$ process), except when this initial value belongs to some set of π measure zero in $\mathbb{R}_0^d \times W$.

Next we aim to show that the π-exceptional set, denoted $\mathcal{E} \subset \mathbb{P}^{d-1} \times W$, over which the P-a.s. limit in (*) may fail, is necessarily of the form $\mathbb{P}^{d-1} \times B$ with $B \subset W$, $\pi_\xi(B) = 0$. This will show that that, if $\xi_0(\omega) \in B^c$, the P-a.s. limit given by Birkhoff's Ergodic Theorem holds for all values of the initial random variable s_0 (and hence does not depend on the initial value in the $\mathbb{R}_0^d$ component of the Markov pair process $\{(x_n, \xi_n)\}$). Since, by stationarity of the $\{\xi_n(\omega)\}$ process, we will then have $P(\omega \in \Omega : \xi_0(\omega) \in B) = \pi_\xi(B) = 0$, this will prove the theorem.

For any $A \in \mathcal{B}(\mathbb{P}^{d-1} \times W)$, let $A_\xi \equiv \{s \in \mathbb{P}^{d-1} : (s, \xi) \in A\}$ and, as in Subsection 4.2, let m_C denote the volume element on $\mathbb{P}^{d-1}$ restricted to unique maximal invariant control set $C \subset \mathbb{P}^{d-1}$ for the dynamical control system $s_{n+1} = f(s_n, u_n)$ associated with $s_{n+1} = f(s_n, \xi_n(\omega))$.

Then $m_C(\mathcal{E}_\xi) + m_C(\mathcal{E}_\xi^c) = m_C(C) = \alpha > 0$.

(Note that, for fixed $\xi \in W$, $\mathcal{E}_\xi^c$ represents the subset of $\mathbb{P}^{d-1}$ over which the P-a.s. limit statement of Birkhoff's Ergodic Theorem holds, while $\mathcal{E}_\xi$ represents the subset of $\mathbb{P}^{d-1}$ over which the P-a.s. limit statement of Birkhoff's Ergodic Theorem does not hold.)

- If $m_C(\mathcal{E}_\xi^c) > 0$, then $\pi(\mathcal{E}_\xi^c \times Q) > 0$ ($Q = \text{supp}(\pi_\xi)$) and, by Theorem 4.2.3 and Remark 4.2.6 (2), there is a stopping time $\tau \equiv \tau((s, \xi), \mathcal{E}_\xi^c \times W)$, $P(\tau < \infty) = 1$, for which

$$P(s_\tau \in \mathcal{E}_\xi^c \mid (s_0, \xi_0) = (s, \xi)) \equiv P((s_\tau, \xi_\tau) \in \mathcal{E}_\xi^c \times W \mid (s_0, \xi_0) = (s, \xi)) = 1,$$

for all $(s, \xi) \in \mathbb{P}^{d-1} \times W$. In particular, this is true for all $s \in \mathcal{E}_\xi$, $\xi \in W$. Therefore, by the strong Markov property, $\mathcal{E}_\xi = \phi$ and $\mathcal{E}_\xi^c = \mathbb{P}^{d-1}$ for all $\xi \in W$.

- If $m_C(\mathcal{E}_\xi) > 0$, the same reasoning as above shows that $\mathcal{E}_\xi^c = \phi$ and $\mathcal{E}_\xi = \mathbb{P}^{d-1}$ for all $\xi \in W$.

Hence, $\mathcal{E} = \bigcup \{\mathcal{E}_\xi ; \xi \in W) = \mathbb{P}^{d-1} \times B$, where $B = \{\xi \in W : \mathcal{E}_\xi = \phi\}$ and, since, by Theorem 4.2.1 (c), $m_C \times \pi_\xi \ll \pi$, $\pi_\xi(B) = 0$. According to the discussion at the begining of this argument, this shows that, for some set $\Omega' \subset \Omega$, $P(\xi_0(\omega) \in \Omega') = 1$, all $\omega \in \Omega'$ and all $x \in \mathbb{R}_0^d$, $\lambda(x) = E_\pi\left[\log\left|A(\xi_0(\omega))\frac{x}{|x|}\right|\right] = \lambda$.

It remains to show that $\lambda = \lim_{n\to\infty} \frac{1}{n} \log \|A(\xi_n(\omega)) \ldots A(\xi_0(\omega))\|$ and that $\lambda = \lambda_1$.

To do this, recall that, because of their equivalence, we can pick any norm for A. We use the usual operator norm, i.e., $\|A(\xi)\| = \sup\{|A(\xi)x| ; |x| = 1\}$. Letting $\{e_1(\omega), \ldots, e_d(\omega)\}$ be a normal basis (whose existence, for all $\omega \in \Omega_0$, is guaranteed by Remark 2.9.1 , but which depends on ω), we have

$$\begin{aligned} |x|\ |A(\xi_n(\omega)) \ldots A(\xi_0(\omega)) \tfrac{x}{|x|}| \\ \leq |x|\ \|A(\xi_n(\omega)) \ldots A(\xi_0(\omega))\| \\ \leq d\,|x| \max_{e_i} |A(\xi_n(\omega)) \ldots A(\xi_0(\omega)) e_i(\omega)|. \end{aligned}$$

So, ω-wise and for $\omega \in \Omega_0 \cap \Omega'$, $P(\xi_0(\omega) \in \Omega_0 \cap \Omega') = 1$,

$$\begin{aligned} \lambda &\leq \overline{\lim_{n\to\infty}} \tfrac{1}{n} \log |x|\ \|A(\xi_n(\omega)) \ldots A(\xi_0(\omega))\| \\ &\leq \overline{\lim_{n\to\infty}} \tfrac{1}{n} \log d\,|x| \max_{e_i} |A(\xi_n(\omega)) \ldots A(\xi_0(\omega)) e_i(\omega)| \\ &= \overline{\lim_{n\to\infty}} \tfrac{1}{n} \max_{e_i} \log |A(\xi_n(\omega)) \ldots A(\xi_0(\omega)) e_i(\omega)| \\ &= \max_{e_i} \left[\overline{\lim_{n\to\infty}} \tfrac{1}{n} \log |A(\xi_n(\omega)) \ldots A(\xi_0(\omega)) e_i(\omega)|\right] \\ &= \max_{e_i} \lambda(e_i(\omega)) = \lambda. \end{aligned}$$

Repeating this argument with $\underline{\lim}$'s shows that

$$\lambda = \lim_{n\to\infty} \tfrac{1}{n} \log \|A(\xi_n(\omega)) \ldots A(\xi_0(\omega))\|.$$

Since, moreover $\{e_1(\omega), \ldots, e_d(\omega)\}$ is a normal basis. i.e.,

$$\lambda\left[\sum_{i=1}^{d} c_i e_i(\omega)\right] = \max(\lambda(e_i(\omega)) ; i \text{ such that } c_i \neq 0),$$

and $\lambda = \max_{e_i} \lambda(e_i(\omega))$, λ must be the maximal Lyapunov exponent, i.e., $\lambda = \lambda_1$.
This proves our last statement. ∎

Remark 1.

Besides the use of Birkhoff's Ergodic Theorem to prove the uniqueness of the Lyapunov exponent (P-a.s.) and hence the independence of this exponent from the value of the initial random variable x_0, one can use another quite illuminating approach based on a pathwise argument.
In the following, we fix $\omega \in \Omega$ and, whenever the statements made only hold for P–almost all ω's, we assume that this fixed ω does not belong to the (finite) union of all the exceptional sets.

Let $\{e_1(\omega), \ldots, e_d(\omega)\}$ be a normal basis for $\mathbb{R}^d_0$ (such a normal basis exists for almost all ω's and depends on ω). Then there must be $i_0(\omega)$ such that $\lambda_1 = \lambda(e_{i_0(\omega)}(\omega))$. Hence, $\lambda(x) < \lambda(e_{i_0(\omega)}(\omega))$ if and only if, for all $y \in S_{x\,\xi(\omega)}$ (recall that $S_{x\,\xi}$ denotes the random orbit of x, with $\xi_0(\omega) = \xi$), we have $y = \sum_{i=1}^{d} \alpha_i(\omega, y)\, e_i(\omega)$ with $\alpha_{i_0}(\omega, y) = 0$, or, in other words, if and only if

$$S_{x\,\xi(\omega)} \subset \text{span}(e_1(\omega), \ldots, e_{i_0(\omega)-1}(\omega), e_{i_0(\omega)+1}(\omega), \ldots, e_d(\omega)).$$

Now, all of our hypotheses imply that, after projection of our system onto $\mathbb{P}^{d-1}$, all open sets in the unique maximal invariant control set $C \subset \mathbb{P}^{d-1}$ are, loosely speaking, "strongly recurrent", i.e., even if $\{\xi_n\}$ is not strongly recurrent, all sets of the form $O \times W$, $O \subset C$ and O open, satisfy $L((s, \xi), O \times W) = 1$) for all $(s, \xi) \in \mathbb{P}^{d-1} \times W$ (see Remark 4.2.6 (2)). Therefore, with probability one (i.e., for P–almost all ω's),

$S_{s\,\xi(\omega)} \cap O \neq \phi$ for all open sets $O \subset \mathbb{P}^{d-1}$.

But span $(e_1(\omega), \ldots, e_{i_0(\omega)-1}(\omega), e_{i_0(\omega)+1}(\omega), \ldots, e_d(\omega))$ is a lower dimensional subspace of $\mathbb{R}_0^d$ and its projection into $\mathbb{P}^{d-1}$, denoted $S(\omega)$, satisfies $P^{d-1} \setminus S(\omega)$ is open. Hence, for P–almost all ω's and for all $s \in \mathbb{P}^{d-1}$, $S_{s\,\xi(\omega)}$ leaves $S(\omega)$. Accordingly, for all $x\,\mathbb{R}_0^d$,

$$S_{x\,\xi(\omega)} \cap \left[\text{span}\,(e_1(\omega), \ldots, e_{i_0(\omega)-1}(\omega), e_{i_0(\omega)+1}(\omega), \ldots, e_d(\omega))\right]^c$$

for P–almost all ω's. Therefore, for all $\omega \notin \tilde{\Omega}$, $P(\tilde{\Omega}) = 0$, and for all $x \in \mathbb{R}_0^d$,

$$\lim_{n\to\infty} \frac{1}{n} \log |x(n, (x_0, \xi_0), \omega)| = \lambda_1.$$

Starting in Section 5, we have assumed that the $\{\xi_n\}$ process is stationary (and ergodic). It then immediately followed that, independently of the value in $\mathbb{R}_0^d$ of the random variable x_0, $\lambda(x) = \lambda_1$ P-a.s. (see Theorem 1). Nevertheless, it is not necessary to assume stationarity of the noise process $\{\xi_n\}$, provided that Assumption 5 from Subsection 4.1 is satisfied. A statement to this effect is the purpose of the following corollary.

Corollary 1.

If the noise process $\{\xi_n\}$ is strongly recurrent (π_ξ-recurrent) and satisfies Assumption 5, then the statement of Theorem 1 holds.

Proof.

Again denote by $\mathcal{E} \subset \mathbb{P}^{d-1} \times W$ the π-null set over which the limit statement of Birkhoff's Ergodic Theorem may fail. Then $\mathcal{E}^c$ has full π measure and since, by Theorem 4.2.3, the pair process $\{(s_n, \xi_n)\}$ is π-recurrent, it follows that there is a

stopping time $\tau \equiv \tau((x, \xi), \mathcal{E}^c)$ such that

$$P((s_\tau, \xi_\tau) \in \mathcal{E}^c \mid (s_0, \xi_0) = (s, \xi)) = 1 \text{ for all } (s, \xi) \in \mathbb{P}^{d-1} \times W.$$

The result then follows immediately from the strong Markov property. ■

Remark 2.

If, still under Assumption 5, the process $\{\xi_n\}$ is neither stationary nor strongly recurrent, there is a π_ξ-null set over which the initial distribution of the $\{\xi_n\}$ process may put some mass and within which the $\{\xi_n\}$ process may remain with positive probability. There could therefore be a positive probability that the path of the $\{\xi_n(\omega)\}$ process will remain in the π-exceptional set in (*), which implies that $\lim_{n\to\infty} \frac{1}{n} \log |x(n, (x, \xi_0(\omega)), \omega)|$ may not be equal to λ_1 with probability one.

Remark 3.

If the $\{\xi_n\}$ process is strongly recurrent (and possibly stationary), then, by Theorem 4.2.3, the pair process $\{(s_n, \xi_n)\}$ on $\mathbb{P}^{d-1} \times W$ is strongly recurrent.

1) If the invariant set $C \times Q$ ($Q = \text{supp}(\pi_\xi)$) is not decomposable into a cycle, the above result states that, with probability one, every solution of the stochastic difference equation $x_{n+1} = A(\xi_n) x_n$ has, in distribution, the asymptotic form

$$x_n \sim \tilde{s}_n e^{\lambda n},$$

where $\{(\tilde{s}_n, \xi_n)\}$ represents the unique stationary solution on $\mathbb{P}^{d-1} \times W$ of the process $\{(s_n, \xi_n)\}$ corresponding to the unique invariant measure π, i.e., $(\tilde{s}_0, \xi_0) \sim \pi$.

Indeed, we know (see Theorem 2.8.3 (a)) that, on $\mathbb{P}^{d-1} \times W$, the Markov chain

$\{(s_n, \xi_n)\} = \{(x_n |x_n|^{-1}, \xi_n)\}$, $x_n \neq 0$, will approach the stationary solution $\{(\tilde{s}_n, \xi_n)\}$ in the sense that, for all $B \in \mathcal{B}^{\infty}(\mathbb{P}^{d-1} \times W)$ (the Borel sets on the path space $(\mathbb{P}^{d-1} \times W)^{\mathbb{N}}$ for the pair process $\{(s_n, \xi_n)\}$) and with ν denoting any initial distribution on $\mathbb{P}^{d-1} \times W$ for the process $\{(s_n, \xi_n)\}$,

$$P_\nu[((s_n, \xi_n), (s_{n+1}, \xi_{n+1}), \ldots) \in B] \underset{n\to\infty}{\longrightarrow} P[((\tilde{s}_n, \xi_n), (\tilde{s}_{n+1}, \xi_{n+1}), \ldots) \in B],$$

and hence that $\bar{\mu}((s, \xi), .)$, the kernel of $\{(s_n, \xi_n)\}$, converges weakly to $\pi(.)$. So, $(x_n, \xi_n) = (s_n |x_n|, \xi_n)$ will approach in distribution $(\tilde{s}_n e^{\lambda n}, \xi_n)$, which is our statement.

2) If the invariant set $C \times Q \subset \mathbb{P}^{d-1} \times W$ can be decomposed into a cycle, i.e., if

$$C \times Q = \bigcup_{i=1}^{d} (C_i' \times Q) \quad \text{for some } d > 1$$

(see Remark 4.2.7), then the initial distribution ν of the $\{(s_n, \xi_n)\}$ process may not converge to the distribution of the stationary distribution $\{(\tilde{s}_n, \xi_n)\}$ and we can only write

$$x_n = s_n |x_n| \sim s_n e^{\lambda n},$$

where now this asymptotic form is attained P-a.s. (since $|x_n| \longrightarrow e^{\lambda n}$ P-a.s.).

Remark 4.

1) Note that, comparing the stochastic case with the linear deterministic case $x_{n+1} = A x_n$, we see that the top Lyapunov exponent λ_1 plays the role of the top real part of the eigenvalues of A, while the marginal of π on $\mathbb{P}^{d-1}$ describes the rotational behavior of the system.

2) Finally, recall that exponential stability of the stochastic system will depend on the sign of λ_1, with a negative λ_1 value yielding stability.

5.3. Moment Stability

In order to obtain more insight concerning the behavior of the process (x_n, ξ_n) and, in particular, to investigate the large deviation properties of our system, we need to introduce the notion of moment Lyapunov exponents.

Definition 1.

The p^{th} moment Lyapunov exponent for a cocycle $C(n, \omega)$ at ω in the direction of $x \in \mathbb{R}_0^d$ is defined by

$$g(p, x) = \overline{\lim_{n\to\infty}} \frac{1}{n} \log E_P \, |C(n, \omega)\, x|^p, \quad p \in \mathbb{R}$$

where P is the probability measure on the path space $\Omega = W^{\mathbb{N}}$ generated by the finite dimensional distributions of the $\{\xi_n(\omega)\}$ process.

In the setup that occupies us and with our notation, the cocycle under consideration is of the form $C(n, \omega) = A(\xi_n(\omega))\, A(\xi_{n-1}(\omega)) \ldots A(\xi_0(\omega))$ (see Subsection 5.1) and Definition 1 becomes, with $p \in \mathbb{R}$,

$$g(p, x) = \overline{\lim_{n\to\infty}} \frac{1}{n} \log E_P \, |A(H_n\, \omega)\, A(H_{n-1}\, \omega) \ldots A(H_0\, \omega)\, x|^p,$$

or equivalently,

$$g(p,x) = \overline{\lim_{n\to\infty}} \frac{1}{n} \log E\,|A(\xi_n(\omega)) \ldots A(\xi_0(\omega))\,x|^p,$$

where the expectation is taken over the joint distribution of $(\xi_n(\omega), \ldots, \xi_0(\omega))$.

In preparation for a series of results concerning p^{th} moment Lyapunov exponents, we will strengthen one of our assumptions from Subsection 5.1. To ensure the finiteness, for all $n \in \mathbb{N}$, of the p^{th} moments $E\,|A(\xi_n) \ldots A(\xi_0)\,x|^p$, it is convenient (but not really needed (see Arnold and Kliemann (1986, Remark 1.1))) to further assume (as in Arnold (1984, p. 794)) that, besides Assumptions (a) through (g), we also have:

h) the manifold W is compact.

This assumption will also be used in Proposition 2 (see Remark 2).

Now, the pair process $\{(s_n, \xi_n)\}$ is Feller (see Proposition 4.1.1) and so, the semigroup generating this pair process maps continuous functions to continuous functions (Feller property). Let us set up some additional notation and give a lemma. Consider, on the space of continuous functions from $\mathbb{P}^{d-1} \times W$ to $\mathbb{P}^{d-1} \times W$, $C(\mathbb{P}^{d-1} \times W)$, the family of operators $\{L(p)\,;\, p \in \mathbb{R}\}$ defined by

$$L(p)\,f(s,\xi) = L\,f(s,\xi) + p \log |A(\xi)\,s|,$$

where L is the generator of the Markov chain $\{(s_n, \xi_n)\}$. Note that the multiplication operator $B = p \log |A(\xi)\,s|$ is bounded and that the domains of $L(p)$, $D(L(p))$, and of L, $D(L)$, agree, i.e., $D(L(p)) = D(L)$.

Now, if $\{T_n\,;\, n \in \mathbb{N}\}$ is a semigroup on the Banach space $\mathcal{E} \equiv C(\mathbb{P}^{d-1} \times W)$,

1) T_n strongly continuous means that $\lim_{n \to 0} T_n f = f$ for all $f \in \mathcal{E}$ and

2) T_n irreducible means that, for all $f \geq 0$ $(f \neq 0)$ and all $(s, \xi) \in \mathcal{E}$, there is an $n \in \mathbb{N}$ such that $T_n f(s, \xi) > 0$.

Lemma 1.

For each $p \in \mathbb{R}$, $L(p)$ is the generator of a strongly continuous semigroup $\tilde{T}_n(p)$ defined by

$$\tilde{T}_n(p) f(s, \xi) = E\left[\left[\exp\left[p \sum_{k=0}^{n} \log |A(\xi_k) s_k|\right]\right] f(s_n, \xi_n) \,\middle|\, (s_0, \xi_0) = (s, \xi)\right].$$

In particular,

1) $\tilde{T}_n(p)$ is positive, i.e., maps positive functions to positive functions,

2) $\tilde{T}_n(p)$ is compact for all $p \in \mathbb{R}$ if and only if $\tilde{T}_n(0)$ is, and

3) $\tilde{T}_n(p)$ is irreducible for all $p \in \mathbb{R}$ if and only if $\tilde{T}_n(0)$ is.

Proof.

The definition of $\tilde{T}_n(p)$ comes from the Feynman–Kac formula. The other statements are based on arguments from functional analysis and can be found in Lemma 2.1 of Arnold (1984). ∎

Remark 1.

Note that, using the same calculation as in Theorem 5.2.1,

$$\exp\left[p \sum_{k=0}^{n} \log |A(\xi_k) s_k|\right] = \exp p \log |A(\xi_n) \ldots A(\xi_0) x_0 / |x_0||,$$

$$= |A(\xi_n) \dots A(\xi_0) x_0/|x_0||^p$$

and hence,

$$\tilde{T}_n(p) f(s,\xi) = E\left[\left[\exp\left[p \sum_{k=0}^{n} \log |A(\xi_k) s_k|\right]\right] f(s_n, \xi_n) \,\middle|\, (s_0, \xi_0) = (s, \xi)\right]$$

$$= E\left[|A(\xi_n) \dots A(\xi_0) x_0/|x_0||^p f(s_n, \xi_n) \,\middle|\, (s_0, \xi_0) = (s, \xi)\right].$$

We can now apply the Perron–Frobenius theory.

Proposition 1.

Suppose $\tilde{T}_n(0)$ is compact and irreducible and define $g(p)$ by

$$g(p) = \lim_{n\to\infty} \frac{1}{n} \log \|\tilde{T}_n(p)\|.$$

Then

1) $\exp(n\, g(p))$ is an eigenvalue of $\tilde{T}_n(p)$. It is simple and larger in magnitude than any other eigenvalue.

2) For $f \in C(\mathbb{P}^{d-1} \times W)$, $f \geq 0$, $f \not\equiv 0$, and for all $(s, \xi) \in \mathbb{P}^{d-1} \times W$,

$$\lim_{n\to\infty} \frac{1}{n} \log \tilde{T}_n(p) f(s,\xi) = g(p),$$

uniformly in $\mathbb{P}^{d-1} \times W$.

3) $g(p)$ is real analytic on $\mathbb{R}$.

Proof.

First, the limit above exists because, for each $p \in \mathbb{R}$, $\{\tilde{T}_n(p) : n \geq 0\}$ is a sequence of

compact operators. The rest of the proof can be found in Lemma 2.2 of Arnold (1984).

■

Remark 2.

In his proof of Lemma 1, Arnold (1984) shows that

$$\tilde{T}_n(p) f(s, \xi) = \exp[n g(p)] m(f) q(s, \xi) (1 + o(1)),$$

where $m \in C^*(\mathbb{P}^{d-1} \times W)$ (the dual space of $C(\mathbb{P}^{d-1} \times W)$) is a unique, positive, and finite (since W is now assumed to be compact) measure, and q ($\| q \| = 1$) is the positive and continuous eigenfunction corresponding to the eigenvalue $\exp[n g(p)]$. Also, $o(1)$ is uniform in $\mathbb{P}^{d-1} \times W$, which means that, if $\{k_n ; n \geq 1\} \subset \mathbb{P}^{d-1} \times W$ converges to $k \in \mathbb{P}^{d-1} \times W$ and, $f \in C(\mathbb{P}^{d-1} \times W)$, then the ratio of distances $\frac{\rho(f(k_n), f(k))}{\rho(k_n, k)}$ converges to zero as $n \to \infty$ uniformly in k. Recall also that $m(f) = \int f(x) m(dx)$.

These facts will be used in the following proof.

Proposition 2.

Let $g(p)$ be as in Proposition 1. Then

$$g(p) = g(p, x) \equiv \lim_{n \to \infty} \frac{1}{n} \log E \, |A(\xi_n) \ldots A(\xi_0) x|^p,$$

i.e., the limit $g(p, x)$ exists and does not depend on $x \in \mathbb{R}_0^d$.

Proof.

To prove this result, we will give a slight modification of the proof given in Arnold (1984, Theorem 2.1). First we show that $\tilde{T}_n(0)$ satisfies the assumptions of

Proposition 1.

The compactness of $\tilde{T}_n(0)$ follows immediately from the equation (see Remark 2)

$$\tilde{T}_n(0) f(s,\xi) = E\left[|A(\xi_n)\dots A(\xi_0)\, x_0/|x_0||^0 f(s_n,\xi_n) \mid (s_0,\xi_0) = (s,\xi)\right]$$

$$= E\left[f(s_n,\xi_n) \mid (s_0,\xi_0) = (s,\xi)\right].$$

Indeed, if $\| f \| \leq 1$, the expectation above is certainly bounded and this means compactness.

The irreducibility of $\tilde{T}_n(0)$ is immediate if there exists a time n such that, for all open sets $B \subset \mathbb{P}^{d-1} \times W$ and for all $(s,\xi) \in \mathbb{P}^{d-1} \times W$, $\bar{\mu}(n,(s,\xi),B) > 0$, where $\bar{\mu}(.,.)$ denotes the kernel of the pair process $\{(s_n,\xi_n)\}$. Hence, by Proposition 4.2.1 (or Theorem 4.2.2 and Corollary 4.2.2), $\tilde{T}_n(0)$ is certainly irreducible on $C(C \times Q)$, with $Q = \text{supp}(\pi_\xi)$, π_ξ denoting the initial distribution (and unique invariant probability measure) for the stationary process $\{\xi_n\}$.

For some fixed $x \in \mathbb{R}_0^d$ such that $(s,\xi) \in C \times Q$ with $s = \frac{x}{|x|}$, we then have

$$E\,|A(\xi_n)\dots A(\xi_0)\, s|^p = \int_Q \tilde{T}_n(p)\, 1\,(s,\xi)\, \pi_\xi(d\xi)$$

$$= \int_Q \exp[n\, g(p)]\, m(1)\, q(s,\xi)\,(1 + o(1))\, \pi_\xi(d\xi)$$

$$= \exp[n\, g(p)]\left[m(1)\left[\int_Q q(s,\xi)\, \pi_\xi(d\xi)\right] + o(1)\right],$$

where 1 denotes the map $1(s,\xi) = 1$. Taking the logarithm and the limit as $n \longrightarrow \infty$

shows that, if $s = \frac{x}{|x|} \in C$,

$$g(p) = \lim_{n \to \infty} \frac{1}{n} \log E \, |A(\xi_n) \ldots A(\xi_0) x|^p$$

If $s = \frac{x}{|x|} \notin C$, Proposition 4.2.1 (or Theorem 4.2.2 and Corollary 4.2.2) says that, with positive probability, one does enter $C \times Q$ from any $(s, \xi) \in \mathbb{P}^{d-1} \times W$ in finite time. Let $\tau \equiv \tau((s, \xi), C \times Q)$ denote the first hitting time of the set $C \times Q$, starting at $(s, \xi) \in \mathbb{P}^{d-1} \times W$, by the Markov process $\{(s_n, \xi_n)\}$. Hence, for n large enough, there is a time $t < n$ such that $P(\tau = t) > 0$. To simplify the notation, we write $z_k = (s_k, \xi_k)$ (also, $z^0 = (s, \xi)$) and $F(z_k) = \log |A(\xi_k) s_k|$. We then have

$$E \, |A(\xi_n) \ldots A(\xi_0) s|^p$$

$$= E \left[\exp p \sum_{k=0}^{n} F(z_k) \mid z_0 = z^0 \right]$$

$$= E \left[E \left[\exp p \sum_{k=0}^{t-1} F(z_k) \exp p \sum_{k=t}^{n} F(z_k) \mid (z_0, \ldots, z_t) = (z^0, \ldots, z^t) \right] \Big| \, z_0 = z^0 \right]$$

$$= E \left[\exp p \sum_{k=0}^{t-1} F(z_k) \, E \left[\exp p \sum_{k=t}^{n} F(z_k) \mid (z_0, \ldots, z_t) = (z^0, \ldots, z^t) \right] \Big| \, z_0 = z^0 \right]$$

$$= E \left[E \left[\exp p \sum_{k=t}^{n} F(z_k) \mid (z_0, \ldots, z_t) = (z^0, \ldots, z^t) \right] \Big| \, z_0 = z^0 \right] .$$

$$. \, E \left[\exp p \sum_{k=0}^{t-1} F(z_k) \mid z_0 = z^0 \right]$$

Now note that, since the map $A(.) : W \longrightarrow Gl(d, \mathbb{R})$ is bounded with W compact

(Assumption (h)), $F(z) = \log |A(\xi)\, s|$ satisfies, for some $\alpha < \infty$, $|F(z)| \leq \alpha$ for all $z = (s, \xi) \in \mathbb{P}^{d-1} \times W$. Hence, the last term in this product is bounded below by

$$\exp(-|p|\ t\ \alpha),$$

and we obtain

$$E\,|A(\xi_n) \ldots A(\xi_0)\, s|^p$$

$$\geq E\left[E\left[\exp p \sum_{k=t}^{n} F(z_k) \mid z_t = z^t\right] \,\middle|\, z_0 = z^0\right] .$$

$$. \exp(-|p|\ t\ \alpha)$$

$$\geq \exp(-|p|\ t\ \alpha)$$

$$. \int_{C \times Q} E\left[\exp p \sum_{k=t}^{n} F(z_k) \mid z_t = z\right] \bar{\mu}(z^0, dz),$$

writing $dz = ds \times d\xi$.

Now recall that

$$E\left[\exp p \sum_{k=t}^{n} F(z_k) \mid z_t = z\right] = E\left[\exp p \sum_{k=0}^{n-t} F(z_k) \mid z_0 = z\right] = \tilde{T}_{n-t}(p)\, 1(s, \xi),$$

and that $\tilde{T}_{n-t}(p)\, 1(s, \xi) = \exp\left[(n-t)\, g(p)\right] m(1)\, q(s, \xi)\, [1 + o(1)]$

(see Remark 2).

Since q is a continuous positive function, it is bounded below (say by β) on the compact set $C \times Q \subset \mathbb{P}^{d-1} \times W$ and we get that

$$\int_{C \times Q} E\left[\exp p \sum_{k=t}^{n} F(z_k) \mid z_t = z\right] \bar{\mu}(z^0, dz)$$

$$\geq \beta \exp\left[(n-t)\, g(p)\right] m(1)\,[1 + o(1)]\, P(z_t \in C \times Q \mid z_0 = z^0).$$

From this it follows that

$$E\,|A(\xi_n) \ldots A(\xi_0)\, s|^p$$

$$\geq \beta \exp(-|p|\, t\, \alpha)$$

$$\cdot \exp\left[(n-t)\, g(p)\right] m(1)\,[1 + o(1)]\, P(z_t \in C \times Q \mid z_0 = z^0)$$

and, upon dividing by n, taking the logarithm, and letting $n \to \infty$, we conclude that for $x = s\,|x|$ such that $s \notin C$, $g(p, x) \geq g(p)$.

To finish up the proof it then suffices to show that $g(p, x) \leq g(p)$.
In order to do this, let $C' \equiv \{x \in \mathbb{R}_0^d : \frac{x}{|x|} \in C\}$. By Assumption (g), int $C \neq \phi$ and hence int $C' \neq \phi$. Therefore, we can select a basis $\{e_1, \ldots, e_d\} \subset C'$ and write, for all $x \notin C'$, $x = \sum_{i=1}^{d} c_i(x)\, e_i$. But then, since, for fixed $p \in \mathbb{R}$, $g(p, x)$ is a Lyapunov exponent, the fourth property of Lyapunov exponents listed on p. 64 yields that

$$g(p, x) \leq \max\{g(p, e_i)\,;\, 1 \leq i \leq d\} = g(p).$$

■

Corollary 1.

1) $g(0, x) = g(0) = 0$,

2) $g(p) \leq p \log K$, where $K = \max_{\xi \in Q} \|A(\xi)\|$ and $Q = \text{supp}(\pi_\xi)$, and

3) $g(p) \geq p\,\lambda_1$.

Proof.

1) Trivial in view of the definition of $g(p)$.

2) $$\lim_{n\to\infty} \frac{1}{n} \log E\,|A(\xi_n) \ldots A(\xi_0)\,x|^p$$

$$\leq \overline{\lim_{n\to\infty}} \frac{1}{n} \log |x|^p\, E\left[\|A(\xi_n)\|^p \ldots \|A(\xi_0)\|^p\right]$$

$$\leq \overline{\lim_{n\to\infty}} \frac{1}{n} \log E\left[\max_{\xi\in Q} \|A(\xi)\|^{np}\right] = \lim_{n\to\infty} \frac{1}{n} \log K^{np} = p \log K.$$

3) Using the fact that log is concave, we get

$$g(p) = \lim_{n\to\infty} \frac{1}{n} \log E\,\|A(\xi_n) \ldots A(\xi_0)\|^p \geq \lim_{n\to\infty} \frac{1}{n} E \log \|A(\xi_n) \ldots A(\xi_0)\|^p$$

Then, Fatou's lemma yields $g(p) \geq p\,E\,\lambda_1 = p\,\lambda_1$. ■

Proposition 3.

a) The mapping $p \mapsto g(p)$ is analytic, $g'(0) = \lambda_1$, and

b) $g(p)$ is convex.

Proof.

a) Since we know from Proposition 1 that $g(p)$ is analytic, it is enough to show that the left and right derivatives at 0, $g^-(0)$ and $g^+(0)$ respectively, satisfy

$$g^-(0) \leq \lambda_1 \leq g^+(0).$$

By Jensen's inequality, using the concavity of log (and the convexity of $-\log$), we

have, with $p > 0$,

$$g(p) = \lim_{n\to\infty} \frac{1}{n} \log E \|A(\xi_n) \dots A(\xi_0)\|^p \geq \lim_{n\to\infty} \frac{p}{n} E \left[\log \|A(\xi_n) \dots A(\xi_0)\|\right]$$

and

$$-g(-p) = \lim_{n\to\infty} \frac{1}{n} -\log E \|A(\xi_n) \dots A(\xi_0)\|^{-p}$$

$$\leq \lim_{n\to\infty} \frac{p}{n} E \left[\log \|A(\xi_n) \dots A(\xi_0)\|\right].$$

Hence

$$g^+(0) = \lim_{p\downarrow 0} \frac{g(p)}{p} \geq \lim_{p\downarrow 0} \lim_{n\to\infty} \frac{1}{n} E \left[\log \|A(\xi_n) \dots A(\xi_0)\|\right] = \lambda_1,$$

while, from the left, we get

$$g^-(0) = \lim_{p\downarrow 0} \frac{g(-p)}{-p} \leq \lim_{p\downarrow 0} \lim_{n\to\infty} \frac{1}{n} E \left[\log \|A(\xi_n) \dots A(\xi_0)\|\right] = \lambda_1.$$

b) Convexity means that, for all $\alpha \in [0,1]$,

$$g(\alpha p_1 + (1-\alpha) p_2) \leq \alpha g(p_1) + (1-\alpha) g(p_2).$$

Now,

$$\alpha g(p_1) + (1-\alpha) g(p_2)$$

$$= \lim_{n\to\infty} \frac{1}{n} \left[\alpha \log E \|A(\xi_n) \dots A(\xi_0)\|^{p_1} + (1-\alpha) \log E \|A(\xi_n) \dots A(\xi_0)\|^{p_2}\right]$$

$$= \lim_{n\to\infty} \frac{1}{n} \log \left[\left[E \|A(\xi_n) \dots A(\xi_0)\|^{\frac{\alpha p_1}{\alpha}}\right]^{\alpha} \cdot \left[E \|A(\xi_n) \dots A(\xi_0)\|^{\frac{(1-\alpha)p_2}{1-\alpha}}\right]^{1-\alpha}\right]$$

$$\geq \lim_{n\to\infty} \frac{1}{n} \log E \|A(\xi_n) \dots A(\xi_0)\|^{\alpha p_1 + (1-\alpha) p_2} = g(\alpha p_1 + (1-\alpha) p_2).$$

The inequality is a consequence of Minkowski's (Hölder's) inequality

$$E\,|XY| \le \left[E\,|X|^p\right]^{1/p} \cdot \left[E\,|X|^q\right]^{1/q} \quad \text{for } 1 < p < \infty \text{ and } \frac{1}{p} + \frac{1}{q} = 1 \text{ with } \alpha = \frac{1}{p}$$

and $(1-\alpha) = \frac{1}{q}$.

■

Proposition 4.

The mapping $p \mapsto \frac{g(p)}{p} = \frac{g(p,x)}{p}$ is increasing.

Proof.

By Lyapunov's inequality (see, e.g., Chung (1974, p. 47)), we have

$$\left[E\,|X|^r\right]^{1/r} \le \left[E\,|X|^s\right]^{1/s} \qquad \text{for } 0 < r < s.$$

So, $\frac{1}{r}\log E\,|X|^r \le \frac{1}{s}\log E\,|X|^s$, and the result follows for $p \ge 0$.

For $p < 0$, start with, for $s < r < 0$,

$$\left[E\,\frac{1}{|X|^{-r}}\right]^{-1/r} \le \left[E\,\frac{1}{|X|^{-s}}\right]^{-1/s} \quad \text{to get } \frac{1}{s}\log E\,|X|^s \le \frac{1}{r}\log E\,|X|^r.$$

For $p = 0$, the result follows from the above and analyticity.

■

The result of Proposition 3 above is the usual relationship between sample and p^{th} moment Lyapunov exponents which was already proved by various authors under several different setups (see, for example, Arnold (1984), Arnold et al. (1986a), and Arnold et al. (1986b)). The relation between λ_1 and g (p) is " tight" for small p (i.e., $\lambda_1 < 0$ implies g (p) < 0 for small p) but, in the continuous time case, it is well known from the above authors that one can have $\lambda_1 < 0$ but g (p) > 0 for large p. Quoting Arnold et al. (1986a, Remark 5.2), this would be an indication of the

following phenomenon: Although, with probability one,

$$x(n,(x,\xi_0(\omega)),\omega) = A(\xi_n(\omega)) \ldots A(\xi_0(\omega))\, x \longrightarrow 0 \text{ as } n \longrightarrow \infty,$$

there are, even for large n, a few paths with small probability for which the norm $|x(n,(x,\xi_0(\omega)),\omega)|$ is still large enough to make $E\,|x(n,(x,\xi_0(\omega)),\omega)|^p$ large and cause p^{th} mean instability, i.e., an exponential growth rate $g(p) > 0$. In the discrete time case, the possibility of a similar behaviour is to be expected (take a continuous time system with $\lambda_1 < 0$ but $g(p) > 0$ for some large p, discretize it, but use very small discrete time increments ; such a discrete time system will exhibit a growth behavior similar to the one of the continuous time system from which it arose). Arnold et al. (1986a) studied this type of behavior in the case of stochastic differential equations and we will follow their approach to investigate the discrete time case which occupies us.

Define the function $\gamma(p)$ by

$$\gamma(p) = \begin{cases} \lambda_1, & p = 0 \\ \dfrac{g(p)}{p}, & p \neq 0 \end{cases}$$

By Propositions 3 and 4 and Corollary 1, $\gamma(.)$ is analytic, increasing, and bounded. Hence the limits as $p \longrightarrow \pm\infty$ exists and we may define

$$\gamma^+ = \lim_{p \to \infty} \gamma(p)$$

$$\gamma^- = \lim_{p \to -\infty} \gamma(p).$$

First we describe the basic features of the function $\gamma(.)$ in the form of two propositions.

Proposition 5.

Either $g(p) = \lambda_1 p$ for all $p \in \mathbb{R}$ and hence $\gamma(p) \equiv \lambda_1$ or $\gamma(p)$ is strictly increasing.

Proof.

We need to show that, if $\gamma(.)$ is not strictly increasing, then it must be that $\gamma(p)$ is constant $(= \lambda_1)$ for all p.

If γ is not strictly increasing, then γ is constant on some open set of p values. But, by analyticity, this automatically implies that $\gamma(p)$ is constant and hence that $\gamma(p) = \lambda_1$, i.e., that $g(p) = \lambda_1 p$.

■

Remark 3.

1) If $\lambda_1 < 0$, three cases are possible:

   (a) If $g(p)$ is not strictly convex, then $g(p) = \lambda_1 p$ for all $p \in \mathbb{R}$.

   If $g(p)$ is strictly convex, then

   (b) $g(p) < 0$ for all $p < 0$, or

   (c) $g(p_0) = 0$ for some $p_0 > 0$.

2) If $\lambda_1 > 0$, three cases are also possible:

   (a) If $g(p)$ is not strictly convex, then $g(p) = \lambda_1 p$ for all p.

   If $g(p)$ is strictly convex, then

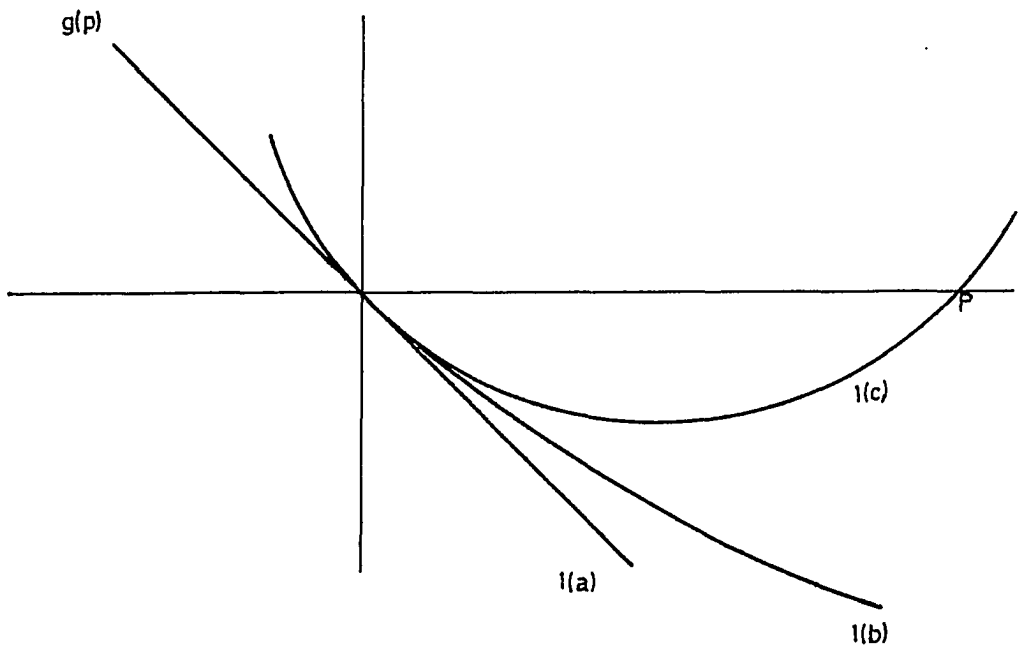
   (b) $g(p) < 0$ for all $p < 0$, or

   (c) $g(p_0) = 0$ for some $p_0 < 0$.

3) If $\lambda_1 = 0$, $g(p) = \lambda_1 p = 0$ for all p or $g(p) > 0$ for all $p \neq 0$.

Proof.

A formal argument is omitted. These assertions simply arise from the picture (given below) of the possible choices for the function g (p):



Remark 4.

Under $\lambda_1 < 0$, Case (a) of the Remark 3 implies that $\gamma^+ = \lambda_1 < 0$, Case (b) that $\gamma^+ \leq 0$, and Case (c) that $\gamma^+ > 0$. So, $\gamma^+ \leq 0$ if and only if $g(p) < 0$ for all $p > 0$, i.e., if and only if we have stable moment Lyapunov exponents for all $p > 0$. Moreover, still under $\lambda_1 < 0$, we see that even if $\gamma^+ > 0$, $g(p) < 0$ for $p \in (0, p_0)$, i.e., that sample stability implies moment stability at least for small p values. Therefore, it appears that the constant γ^+ is of similar importance as λ_1 and contains valuable information about the fine structure of the system, i.e., about path behaviors which are not described by λ_1 (see the discussion following Proposition 4).

5.4. Large Deviations

This subsection will be devoted to a few results concerning the large deviation properties associated with the unique (P-a.s.) Lyapunov exponent of our system. In this, we will essentially follow the discussion on level–1 large deviations found in Arnold and Kliemann (1986). Background information relative to the theory of large deviations can be found in Ellis (1985). The setup is the one described in Subsections 5.2 and 5.3.

Since the p^{th} moment Lyapunov exponent $g(p)$ can be identified to be the free energy of the sequence $\left\{\log \left| A(\xi_n) \ldots A(\xi_0) \frac{x}{|x|} \right| ; n \in \mathbb{N}\right\}$ (see the proof of Theorem 1 for a formal argument), the natural candidate for the level–1 entropy function is the Legendre–Fenchel transform of $g(p)$:

$$I(r) = \sup_{p \in \mathbb{R}} (r\,p - g(p)), \quad r \in \mathbb{R}.$$

In the following, we will systematically write λ for the (P-a.s.) top Lyapunov exponent λ_1. We then have a basic result characterizing I (r).

Proposition 1.

1) If $\gamma^+ = \gamma^-$, then

$$I(r) = \begin{cases} 0, & r = \lambda, \\ \infty, & r \neq \lambda. \end{cases}$$

2) If $\gamma^+ > \gamma^-$, then the following holds:

a)

$$I(r) = \begin{cases} \text{finite}, & r \in (\gamma^-, \gamma^+) \\ \infty, & r \notin [\gamma^-, \gamma^+] \end{cases}$$

b) I is strictly convex and analytic on (γ^-, γ^+).

c) $I(r) \geq 0$, $I(r) = 0$ if and only if $r = \lambda$, and $I'(\lambda) = 0$.

d) $I''(\lambda) = (g''(0))^{-1}$, i.e.,

$$I(r) = (r-\lambda)^2 (2\,g''(0))^{-1} + o\left[(r-\lambda)^3\right] \quad \text{for } |r-\lambda| \text{ small.}$$

e) I is strictly decreasing on (γ^-, λ) and strictly increasing on (λ, γ^+).

Proof.

1) Since $\lambda \in [\gamma^-, \gamma^+]$, we have $\gamma^- = \gamma^+ = \lambda$. Hence, it must be that $g(p) = \lambda\,p$, and the result follows.

2) a) Rewrite $I(r)$ as $I(r) = \sup_{p \in \mathbb{R}} p(r - \gamma(p))$.

Let $r \in (\gamma^-, \gamma^+)$. Then, since $\gamma(p) \to \gamma^+$ as $p \to \infty$, for large p, $p(r - \gamma(p)) < 0$ while $\gamma(p) \to \gamma^-$ as $p \to -\infty$ implies that, for large negative p, we also have $p(r - \gamma(p)) < 0$. But, for $p = 0$, we have $p(r - \gamma(p)) = 0$. Therefore $\sup_{p \in \mathbb{R}} p(r - \gamma(p)) \geq 0$ and this supremum will be attained for $p \in [p_1, p_2]$, with $-\infty < p_1 < p_2 < \infty$. This supremum is then clearly finite. When $r \notin [\gamma^-, \gamma^+]$, a similar reasoning shows that the supremum will be infinite. For $r \in \{\gamma^-, \gamma^+\}$, i.e., at the discontinuity points, no conclusion can be drawn without a more precise expression for $g(p)$.

2) b) Since $g(p)$ is itself the Legendre–Fenchel transform of $I(r)$ (Theorem 6.4.1 in Ellis (1985)), Theorem 6.5.6 in Ellis (1985) and the analyticity of g imply that $I(r)$ must be strictly convex. Moreover, the analyticity of $g(p)$ also implies the analyticity of $I(r)$ via the equation

$$I(r) = p(g'(r))^{-1} - g((g'(r))^{-1}),$$

(see Rockafellar (1970, Theorem 26.6, p. 259)).

2) c) $I(r) \geq 0$ follows from the argument in the proof of part (2) (a). The other two statements come from the equations

i) $rp = g(p) + I(r)$ if and only if $r = g'(p)$ and

ii) $rp = g(p) + I(r)$ if and only if $p = I'(r)$

(Theorem 6.4.1 in Ellis (1985), equation (ii) being obtained by looking at $g(p)$ as the Legendre–Fenchel transform of $I(r)$).

Taking $p = 0$ (and therefore $r = g'(0) = \lambda$) in (i), we get the equation $0 = g(0) + I(\lambda) = I(\lambda)$. Then, strict convexity of $I(r)$ (see part 2 (b))

ensures that λ is the only root of the function I.

Similarly, taking $r = \lambda$ in (ii) (and hence $p = I'(\lambda)$), gives the equation $\lambda I'(\lambda) = g(I'(\lambda))$. But this equation holds for $p = I'(\lambda) = 0$ and, since, by the strict convexity of I (r), I' is strictly increasing, this solution is unique.

2) d) This result follows from differentiating $g'(p)\,p = g(p) + I(g'(p))$ (see equation (i) above) and from the Taylor expansion of I (r) around λ.

2) e) Since I is strictly convex and assumes its minimum at $r = \lambda$, this statement is immediate.

■

We may then state the following large deviation property:

Theorem 1.

For all $x \in \mathbb{R}_0^d$, we have:

a) $\overline{\lim_{n\to\infty}} \frac{1}{n} \log P\left(\frac{1}{n} \log \left| A(\xi_n) \ldots A(\xi_n) \frac{x}{|x|} \right| \in F\right) \leq -\inf_{r\in F} I(r)$

for each closed set $F \subset \mathbb{R}$, and

b) $\underline{\lim_{n\to\infty}} \frac{1}{n} \log P\left(\frac{1}{n} \log \left| A(\xi_n) \ldots A(\xi_n) \frac{x}{|x|} \right| \in G\right) \geq -\inf_{r\in G} I(r)$

for each open set $G \subset \mathbb{R}$.

Proof.

Define $W_n = \log \left| A(\xi_n) \ldots A(\xi_n) \frac{x}{|x|} \right|$ and let $c_n(p)$ be defined by

$$c_n(p) = \frac{1}{n} \log E\left[\exp\left(p \log \left| A(\xi_n) \ldots A(\xi_n) \frac{x}{|x|} \right|\right)\right]$$

$$= \frac{1}{n} \log E \left| A(\xi_n) \ldots A(\xi_n) \frac{x}{|x|} \right|^p.$$

Then $c(p) \equiv \lim_{n\to\infty} c_n(p) = g(p)$ and $c_n(p)$, $n \in \mathbb{N}$, are finite for all $p \in \mathbb{R}$. Moreover, $c(p)$ (which is called the free energy of the sequence $\{W_n\}$) is differentiable for all $p \in \mathbb{R}$.

Therefore, Theorem 2.6.1 in Ellis (1985) applies, which proves our result.

■

Remark 1.

As noted in Arnold and Kliemann (1986, Remark 3.3), Theorem 1 above holds uniformly in the sense that $P(.)$ in (a) and (b) above can be replaced by $\sup_{x/|x|\in C} P(.)$ and $\inf_{x/|x|\in C} P(.)$, respectively. This follows from the fact that $g(p)$ is a uniform limit in C (see Proposition 5.3.1).

Corollary 1.

For each $x \in \mathbb{R}_0^d$ and each $\epsilon > 0$,

$$\lim_{n\to\infty} \frac{1}{n} \log P\left(\left|\frac{1}{n}\log\left|A(\xi_n)\ldots A(\xi_n)\frac{x}{|x|}\right| - \lambda\right| \geq \epsilon\right) = k(\epsilon),$$

where $k(\epsilon) = -\min(I(\epsilon+\lambda), I(\epsilon-\lambda)) < 0$.

Proof.

Let $B(\lambda, \epsilon)$ denote the open ball centered at λ and with radius ϵ. By Theorem 1 (a), we have:

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log P\left(\frac{1}{n}\log\left|A(\xi_n)\ldots A(\xi_n)\frac{x}{|x|}\right| \in B(\lambda,\epsilon)^c\right)$$

$$\leq -\inf\{I(r) ; r \in B(\lambda,\epsilon)^c\} = -\min(I(\lambda+\epsilon), I(\lambda-\epsilon)), \qquad (*1)$$

where the last equality holds because of Proposition 1 (2e).

On the other end, by Theorem 1 (b),

$$\underline{\lim}_{n\to\infty} \frac{1}{n} \log P\left(\frac{1}{n} \log \left| A(\xi_n) \ldots A(\xi_n) \frac{x}{|x|} \right| \in B(\lambda, \epsilon)^c\right)$$

$$\geq \underline{\lim}_{n\to\infty} \frac{1}{n} \log P\left(\frac{1}{n} \log \left| A(\xi_n) \ldots A(\xi_n) \frac{x}{|x|} \right| \in \overline{B(\lambda, \epsilon)}^c\right)$$

$$\geq -\inf \{I(r) ; r \in \overline{B(\lambda, \epsilon)}^c\} = -\min (I(\lambda+\epsilon), I(\lambda-\epsilon)), \qquad (*2)$$

Combining (*1) and (*2) then gives the result.

■

Corollary 2.

Let $\{a_n ; n \in \mathbb{N}\}$ be a sequence of positive real numbers such that $\lim_{n\to\infty} \frac{1}{n} \log a_n = 0$.

a) If $\gamma^- = \gamma^+ = \lambda$, then, for all $x \in \mathbb{R}_0^d$,

$$\lim_{n\to\infty} \frac{1}{n} \log P\left(\frac{1}{n} \log \left| A(\xi_n) \ldots A(\xi_n) \frac{x}{|x|} \right| \geq a_n\right) = \begin{cases} 0 & \lambda > 0 \\ -\infty & \lambda < 0 \end{cases}$$

$$\lim_{n\to\infty} \frac{1}{n} \log P\left(\frac{1}{n} \log \left| A(\xi_n) \ldots A(\xi_n) \frac{x}{|x|} \right| \leq a_n\right) = \begin{cases} 0 & \lambda < 0 \\ -\infty & \lambda > 0 \end{cases}$$

b) If $\gamma^- < \gamma^+$, then, for all $x \in \mathbb{R}_0^d$,

$$\lim_{n\to\infty} \frac{1}{n} \log P\left(\frac{1}{n} \log \left| A(\xi_n) \ldots A(\xi_n) \frac{x}{|x|} \right| \geq a_n\right) = \begin{cases} 0 & \lambda > 0 \\ -I(0) & \lambda < 0 \text{ and } \gamma^+ \neq 0 \end{cases}$$

$$\lim_{n\to\infty} \frac{1}{n} \log P\left(\frac{1}{n} \log \left| A(\xi_n) \ldots A(\xi_n) \frac{x}{|x|} \right| \leq a_n\right) = \begin{cases} -I(0) & \lambda > 0 \text{ and } \gamma^- \neq 0 \\ 0 & \lambda > 0 \end{cases}$$

Proof.

The proof follows the argument for the continuous time case found in Arnold and Kliemann (1986, Corollary 3.5). ■

Remark 2.

1) Corollary 2 holds in particular for $a_n \equiv c$ for all $n \in \mathbb{N}$.

2) Note that $-I(0) = \sup_{p \in \mathbb{R}} (-g(p)) = \min_{p \in \mathbb{R}} g(p)$.

6. LINEAR OSCILLATOR

6.1. Setup

In this section we will discuss the results from a computer simulation based on a discretized version of the linear oscillator with a damping force denoted $\beta \in \mathbb{R}$ and a restoring force denoted $1 + \alpha$, $\alpha \in \mathbb{R}$.

In the continuous time case, the equation of this oscillator is given by the second order differential equation

$$\ddot{z} + 2\beta \dot{z} + (1+\alpha)\, z = 0.$$

Defining $x = \begin{bmatrix} z \\ \dot{z} \end{bmatrix}$, this equation can be rewritten as the two dimensional first order linear equation

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ -1-\alpha & -2\beta \end{bmatrix} x.$$

Since accuracy in approximating the continuous solution of this equation is not the object here, we will discretize it by use of Euler's method, i.e., with $x(t)$ denoting a solution of the above differential equation, by use of only the first derivative in the Taylor expansion of $x(t+h)$ around $x(t)$:

$$x(t+h) = x(t) + h\,\dot{x}(t) + \frac{1}{2}h^2\,\ddot{x}(t) + \ldots$$

Hence, neglecting the terms of order smaller or equal to h^2, we get that

$$x([n+1]\,h) \approx x(nh) + h\,\dot{x}(nh),$$

meaning that

$$\frac{1}{h}[x([n+1]\,h) - x(nh)] \approx \dot{x}(nh).$$

Therefore, if we write $y_{n+1} \equiv [x([n+1]\,h) - x(nh)]$, y_{n+1} represents the increment of the solution $x(t)$ when one starts at $x = x(nh)$ and proceeds for a time h. This gives the difference equation

$$\frac{1}{h} y_{n+1} = \begin{bmatrix} 0 & 1 \\ -1-\alpha & -2\beta \end{bmatrix} x(nh),$$

or

$$\frac{1}{h} x(nh + h) = \frac{1}{h} x(nh) + \begin{bmatrix} 0 & 1 \\ -1-\alpha & -2\beta \end{bmatrix} x(nh)$$

Assuming $h = 1$ for notational simplicity, this yields the equation

$$x(n + 1) = \begin{bmatrix} 1 & 1 \\ -1-\alpha & 1-2\beta \end{bmatrix} x(n).$$

For a general value of the increment h, we get

$$x([n + 1]\,h) = \begin{bmatrix} 1 & h \\ -h\,(1+\alpha) & 1-2h\beta \end{bmatrix} x(nh).$$

This difference equation is still deterministic. But, if we assume that the coefficient α in the restoring force is disturbed, at each step, by a random variable $\theta_n \equiv \cos \xi_n$, where $\{\xi_n\,;\, n \in \mathbb{N}\}$ is a sequence of iid random variables uniformly distributed on the circle $\mathbb{S}^1$, we end up with the now stochastic difference equation

$$x([n + 1]\,h) = \begin{bmatrix} 1 & h \\ -h\,(1+\alpha\,\theta_n) & 1-2h\beta \end{bmatrix} x(nh) \equiv A(h, \xi_n)\, x(nh).$$

Moreover, we can take the initial value $x_0 = x(0)$ to be an $\mathbb{R}^2_o$ valued random variable independent of $\{\xi_n ; n \in \mathbb{N}\}$.

Remark 1.

The choice of a uniformly distributed noise is somewhat arbitrary. Note nevertheless that this choice implies that the random component of the restoring force $(\alpha\, \theta_n)$ is distributed on $[1-\alpha, 1+\alpha]$ (not uniformly) and that the average restoring force will be one. This is in agreement with the condition imposed on the real noise in the continuous case by Arnold and Kliemann (1986). The simulation results given in Subsection 6.3 will then show that averaging the noise to handle such a situation (which in our setup simply corresponds to using the deterministic discretized linear oscillator with unit restoring force) may lead to erroneous conclusions about the stability behavior of the stochastic system. Moreover, the symmetry of this setup allows the assumption $\alpha \geq 0$.

Remark 2.

The choice of an iid sequence to play the role of the Markov chain $\{\xi_n\}$ we dealt with throughout this work is obviously a simplification. Nevertheless, even in this case, the results will prove interesting enough to make the simulation nontrivial.

6.2. Checking the Assumptions

Before using the setup described in Subsection 6.1 for a computer simulation, it is necessary to verify that this stochastic version of the discretized linear oscillator

satisfies the working assumptions of Section 5. This is the purpose of this subsection.

That the initial value x_0 is an $\mathbb{R}_0^2$ valued random variable independent of the process $\{\xi_n ; n \in \mathbb{N}\}$ (Assumption (c)) is part of the setup in Subsection 5.1.

Concerning the other assumptions on the general setup, it is clear that we are working on manifolds ($M = \mathbb{R}_0^2$ and $W = \mathbf{S}^1$) (part of Assumption (a)), and that, for fixed α and β, A (h, .) is a bounded and C^1 map. But the mapping A (h, .) need not map $\mathbb{R}$ into Gl (d, $\mathbb{R}$). Nevertheless, this will be the case if

$$\det (A) = 1 - 2h\beta + h^2 (1 - \alpha\theta) \neq 0 \text{ for all } \theta \in [-1, 1].$$

This will be the case if (and only if) one of the following two conditions holds:

(1) $1 - 2h\beta + h^2 (1 + \alpha) < 0$ or

(2) $1 - 2h\beta + h^2 (1 - \alpha) > 0$.

From (2), it is obvious that this can be achieved by choosing the step size h small enough. Under such a condition imposed on h, Assumption (b) is satisfied.

Nevertheless, for simulation purposes, we need to find an explicit upper bound for h. In order to achieve this, we use (2) and consider several possible cases, corresponding to various choices for α and β.

• Suppose that $\beta < 0$.

If $1 - \alpha \geq 0$, then any choice for h will do.

If $1 - \alpha < 0$, then we need to consider the roots of $1 - 2h\beta + h^2 (1 - \alpha) = 0$ (as a function of h).

These are $\rho = \frac{\beta - (\beta^2 - (1-\alpha))^{\frac{1}{2}}}{1-\alpha}$ and $\rho' = \frac{\beta + (\beta^2 - (1-\alpha))^{\frac{1}{2}}}{1-\alpha}$, with, in this case, $\rho' < 0 < \rho$.

In order to ensure that $1 - 2h\beta + h^2(1-\alpha) > 0$, we must then have $\rho' < 0 < h < \rho$, i.e., since $h > 0$,

$$h < \rho = \frac{\beta - (\beta^2 - (1-\alpha))^{\frac{1}{2}}}{1-\alpha}.$$

- Suppose that $\beta \geq 0$.

  If $\alpha = 1$, then we need $h < (2\beta)^{-1}$ (if $\beta \neq 0$). If $\beta = 0$, any h will do.

  If $\alpha \neq 1$, we need to distinguish two cases:

  - If $\beta^2 - (1-\alpha) < 0$, we have conjugate complex roots. But, $\beta^2 - (1-\alpha) < 0$ implies that $1 - \alpha > 0$ and therefore that $1 - 2h\beta + h^2(1-\alpha) > 0$ for all $h > 0$.
  - If $\beta^2 - (1-\alpha) \geq 0$, we again have to look at the roots ρ and ρ', but we have to consider two further cases:
    - If $1 - \alpha > 0$, we get $0 < \rho \leq \rho'$.

      Then, to ensure $1 - 2h\beta + h^2(1-\alpha) > 0$, we need $0 < h < \rho$ (or $\rho' < h$, but this is not useful). Hence we again get the condition

      $$h < \rho = \frac{\beta - (\beta^2 - (1-\alpha))^{\frac{1}{2}}}{1-\alpha}.$$

    - If $1 - \alpha < 0$, we have $\rho' < 0 < \rho$ and we need $\rho' < 0 < h < \rho$. Once more, this yields the condition

      $$h < \rho = \frac{\beta - (\beta^2 - (1-\alpha))^{\frac{1}{2}}}{1-\alpha}.$$

Therefore, A (h, .) will map $\mathbb{S}^1$ into Gl (d, $\mathbb{R}$) if

$$0 < h < \min\left\{\frac{1}{2M},\ \frac{\beta - (\beta^2 - (1-\alpha))^{\frac{1}{2}}}{1_A(\alpha,\beta)\,(1-\alpha)},\ \frac{\beta - (\beta^2 - (1-\alpha))^{\frac{1}{2}}}{1_B(\alpha,\beta)\,(1-\alpha)}\right\}$$

where $M \equiv \max\{0, \beta\}$ and $1_A(\alpha,\beta)$ and $1_B(\alpha,\beta)$ are, respectively, the indicator functions of the sets A and B, defining $A \equiv \{(\alpha,\beta) \in \mathbb{R}^+ \times \mathbb{R} : \beta < 0, 1-\alpha < 0\}$ and $B \equiv \{(\alpha,\beta) \in \mathbb{R}^+ \times \mathbb{R} : \beta \geq 0, \beta^2 - (1-\alpha) < 0\}$. Here it is understood that, if a term in this minimum is infinite (i.e., if $M = 0$, $1_A(\alpha,\beta) = 0$, $1_B(\alpha,\beta) = 0$, or $\alpha = 1$) then this term does not contribute to the minimum. Also note that, even though the term $\frac{1}{2M}$ only arose from the case $\alpha = 1$, it was kept in the expression for the upper bound for h regardless of the α value. This is because this condition will come handy when proving the validity of Assumption (f) below.

The assumptions on the $\{\xi_n\}$ process (specifically, Assumptions (a), (d) and (e) of Subsection 5.1) are trivially satisfied by a sequence of iid random variables with uniform distribution on $\mathbb{S}^1$. In this case, the invariant distribution for the $\{\xi_n\}$ process, π_ξ, has a density $f(\xi) = (2\pi)^{-1}$ for $\xi \in \mathbb{S}^1$.

In order to verify Assumption (f) we could use the results of Theorem 3.3.2, i.e., we could check that $\dim \Gamma^+\left[\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]\right] = 2$ for all $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in \mathbb{R}^2_0$, where

$$\Gamma = \{\mathrm{Ad}_{\xi_k \cdots \xi_1} X_{\xi_0} ; k \geq 0, \xi_0, \ldots, \xi_k \in \mathbb{S}^1\}.$$

Indeed, if $\operatorname{int} Sz \neq \phi$ for all $z = \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in \mathbb{R}^2_0$, $\operatorname{int} Ss \neq \phi$ certainly holds for all $s \in \mathbb{P}^1$. The computations involved are quite simple but somewhat annoying. Hence, since this simple (linear) situation allows it, we will prove the validity of Assumption (f) by direct computations, without using the above result.

It suffices to compute the product

$$\begin{bmatrix} 1 & h \\ \epsilon_2 & 1-2h\beta \end{bmatrix} \begin{bmatrix} 1 & h \\ \epsilon_1 & 1-2h\beta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\,(1+h\epsilon_1) + y\ 2h\ (1-h\beta) \\ x\,(\epsilon_2 + \epsilon_1\,[1-2h\beta]) + y\ (h\epsilon_2 + [1-2h\beta]^2) \end{bmatrix},$$

where $\epsilon_i = -h\,(1 + \alpha\theta_i)$ may take any value in the open set $(-h\,[1+\alpha], -h\,[1-\alpha])$. Then, if $x \neq 0$, we immediately see that the interior of the positive orbit of any point in $\mathbb{R}_0^2$ (and hence in $\mathbb{P}^1$) must be nonempty. If $x = 0$ (and hence $y \neq 0$), we need to compute one more step:

$$\begin{bmatrix} 1 & h \\ \epsilon_3 & 1-2h\beta \end{bmatrix} \begin{bmatrix} y\ 2h\ (1-h\beta) \\ y\,(h\epsilon_2 + [1-2h\beta]^2) \end{bmatrix}$$

$$= \begin{bmatrix} yh\,(2\ [1-h\beta] + h\epsilon_2 + [1-2h\beta]^2) \\ y\,(2h\epsilon_3\ [1-h\beta] + [1-2h\beta]\ [h\epsilon_2 + [1-2h\beta]^2) \end{bmatrix},$$

which, under the previously given condition $h < \frac{1}{2\beta}$ for $\beta > 0$ (so that $1 - h\beta \neq 0$), enables us to conclude the same result.

It remains to verify the general nondegeneracy Assumption (g). But, from the combination of Theorem 4.1.1, Proposition 4.1.1 and the above result that int $Sx \neq \phi$ for all $x \in \mathbb{R}_0^2$ (or int $Ss \neq \phi$ for all $s \in \mathbb{P}^1$), we are able to immediately conclude that Assumption (g) is satisfied in our setup.

Hence, provided that the time increment h satisfies the upper bound

$$h < \min\left\{\frac{1}{2M}, \frac{\beta - (\beta^2 - (1-\alpha))^{\frac{1}{2}}}{1_A(\alpha,\beta)\ (1-\alpha)}, \frac{\beta - (\beta^2 - (1-\alpha))^{\frac{1}{2}}}{1_B(\alpha,\beta)\ (1-\alpha)}\right\}$$

discussed earlier, the setup described in Subsection 6.1 satisfies all the assumptions given in Subsection 5.1 and can therefore be used for the simulation of a stochastic difference equation to which the results of this thesis apply. The results of this

simulation are outlined in the next subsection.

6.3. Simulation Results

The purpose of this simulation was to investigate the stability of the stochastic difference equation

$$x_{n+1} = \begin{bmatrix} 1 & h \\ -h\,(1+\alpha\theta_n) & 1-2h\beta \end{bmatrix} x_n,$$

$\theta_n \equiv \cos \xi_n$ ($\{\xi_n\}$ is a sequence of iid random variables with uniform distribution on the circle S^1), for various α and β combinations.

Setting $\alpha = 0$, this system can be viewed as a deterministic system

$$x_{n+1} = \begin{bmatrix} 1 & h \\ -h & 1-2h\beta \end{bmatrix} x_n = A(h)\, x_n$$

with damping force β and a unit restoring force which will then be disturbed by some iid noise distributed on $[-\alpha, \alpha]$ and with zero mean. It is therefore useful to discuss the stability behavior of this deterministic system before talking about the stochastic setup.

Consider an arbitrary deterministic difference equation of the form

$$x_{n+1} = A\, x_n, \quad A \in Gl(2, \mathbb{R}),\ x_0 \in \mathbb{R}^2_0.$$

Let e_1 and e_2 denote the eigenvalues of A. If e_1 and e_2 are real and $|e_1| \geq |e_2|$, we set $e = e_1$ and define E to be the unit (generalized) eigenvector corresponding to

$e = e_1$. If e_1 and e_2 are two complex conjugate eigenvalues, we set $e = e_1$ and define E to be the complex unit eigenvector corresponding to e_1. The maximal (in this case geometric) growth rate of the above deterministic system is then obtained by starting at the initial value $x_0 = E$. Indeed, we have

$$|x_n| = f(n)\,|e\,E|^n = f(n)\,|e|^n,$$

where $f(n)$ denotes some polynomial growth of degree n (which could appear if the Jordan form of A is not a diagonal matrix) and $|e| = (e\,\bar{e})^{\frac{1}{2}}$ if e is imaginary, $\bar{e}$ denoting the complex conjugate of e.

The top Lyapunov exponent for this deterministic discrete dynamical system is then given by

$$\lambda = \lim_{n\to\infty} \frac{1}{n} \log |x_n| = \log |e|.$$

The system is therefore exponentially stable if $\lambda < 0$ or, equivalently, if $|e| < 1$.

In the specific deterministic case which concerns us, $x_{n+1} = A(h)\,x_n$, the eigenvalues e_1 and e_2 can easily be computed to be

$$(1 - h\beta) \pm h\,(\beta^2 - 1)^{\frac{1}{2}}.$$

Hence, the formula for the top Lyapunov exponent can be computed and used to investigate the stability properties of this dynamical system. But, to do so, we need to consider several cases:

Case 1. $|\beta| < 1$.

In this case, we have complex conjugate eigenvalues and $|e|$ is given by:

$$|e| = \left[(1-h\beta)^2 + h^2(1-\beta^2)\right]^{\frac{1}{2}}$$

$$= \left[h^2 - 2\beta h + 1\right]^{\frac{1}{2}}.$$

Then $\lambda = \log |e| < 0$ (stable system) if and only if $h^2 - 2\beta h + 1 < 1$, i.e., if and only if $\beta > \frac{h}{2} > 0$.

Therefore, the system is unstable for $-1 < \beta \leq 0$ and, for $0 < \beta < 1$, the system is stable whenever $\beta > \frac{h}{2} > 0$. In particular, the system is unstable if $h \geq 2$.

Moreover, λ is a decreasing function of $\beta \in (-1, 1)$.

Case 2: $|\beta| = 1$.

In this case we have a single eigenvalue with multiplicity 2.

Then $|e| = |1 - h\beta|$ and the system is stable ($|e| < 1$) if and only if $0 < h\beta < 2$.

Therefore, if $\beta = -1$, the system is unstable, while, if $\beta = 1$, the system is stable provided $h < 2$.

Case 3: $\beta < -1$.

In this case we have two distinct real eigenvalues and the system is stable if (and only if) $h(\beta^2 - 1)^{\frac{1}{2}} - h\beta < 0$, which is impossible.

From these first three cases, we can already conclude that:

(a) If $\beta \leq 0$, the deterministic system is unstable for all values of h.

(b) If $0 < \beta \leq 1$, the system will be stable if $h < 2\beta$.

Moreover, over the range $0 < \beta \leq 1$, the minimal value for λ will be $\lambda = \log |1 - h|$, achieved at $\beta = 1$. We also get $\lambda = 0$ at $\beta = \frac{h}{2}$.

Next we need to consider the case $\beta > 1$ (two distinct real eigenvalues). We will need to distinguish several subcases (depending on h):

Case 4: $\beta > 1$.

- Subcase (i): $h < 1$ and $1 < \beta < \frac{1}{h}$ (i.e., $1 - h\beta > 0$).

  Then $|e| = 1 - h\beta + h(\beta^2 - 1)^{\frac{1}{2}}$ and the system will be stable if $h(\beta^2 - 1)^{\frac{1}{2}} < h\beta$. Since this is true for all h ($h < 1$), the system is always stable. Moreover, it is easy to check that, in this case, λ is an increasing function of β ($\beta > 1$).

- Subcase (ii): $h < 1$ and $\frac{1}{h} < \beta$ (i.e., $1 - h\beta < 0$).

  Then $|e| = h\beta - 1 + h(\beta^2 - 1)^{\frac{1}{2}}$ and we have stabilility if $h\beta - 2 + h(\beta^2 - 1)^{\frac{1}{2}} < 0$. In other words, the system is stable if and only if $0 < h(\beta^2 - 1)^{\frac{1}{2}} < 2 - h\beta$ or, if and only if $h^2 - 4h\beta + 4 > 0$. This implies that the system is stable if and only if $(\frac{1}{h} <) \beta < (h^2 + 4) / 4h$ (which is certainly a valid inequality for all $h < 1$). Moreover, λ can be easily verified to be a increasing function of $\beta > \frac{1}{h}$.

- Subcase (iii): $h > 1$ and $1 < \beta < \frac{1}{h}$.

  This subcase is impossible.

- Subcase (iv): $h > 1$ and $\frac{1}{h} < \beta$ (i.e., $1 - h\beta < 0$).

  This subcase follows the reasoning of Subcase (ii).

Putting together Cases 1 through 4, the value of the top Lyapunov exponent (as a function of the damping force β) for the deterministic system $x_{n+1} = A(h)\, x_n$ can be described as follows:

a) If $h < 2$, the top Lyapunov exponent λ is positive and decreasing for $-\infty < \beta < \frac{h}{2}$, negative for $\frac{h}{2} < \beta < (h^2 + 4) / 4h$ (first decreasing on $\frac{h}{2} < \beta < 1$ and then increasing

on $1 < \beta < (h^2 + 4) / 4h)$, positive and increasing (giving an unstable system due to overdamping) for $\beta > (h^2 + 4) / 4h$, and $\lambda = 0$ for $\beta = \frac{h}{2}$ and $\beta = (h^2 + 4) / 4h$. Moreover, λ is minimum at $\beta = 1$ ($\lambda = \log |1 - h|$).

b) If $h \geq 2$, λ follows the same decreasing/increasing pattern as in (a) but λ remains nonnegative for all β values (unstable system).

Following these remarks concerning the stability behavior of the deterministic system $x_{n+1} = A(h)\, x_n$, we can now talk about the results of the stochastic simulation. First we need to discuss the choice for the values of the parameters involved.

1) The parameter α

Since the noise process $\{\xi_n ; n \in \mathbb{N}\}$ is a sequence of iid random variables uniformly distributed on $\mathbb{S}^1$ and therefore $\{\theta_n ; n \in \mathbb{N}\}$, $\theta_n = \cos \xi_n$, is a sequence of iid random variables distributed on $[-1, 1]$, the α parameter basically describes the extent of the random disturbance built into the system. During preliminary observations, it was found that, for $\alpha < 1.00$, the stochastic system was stable for all the (attempted) (h, β) combinations which also yielded exponential stability for the deterministic system (even though the stochastic system did show a slower rate of exponential decay than the corresponding deterministic system). On the other hand, large α values impose the use of very small incremental h values (because of the upper bound condition on h). This was not desired since, with h small and for small β values, the (λ, β) curves for the stochastic systems were

found to be very close to the deterministic curve, which was inconvenient for plotting purposes. Therefore α was taken to range from 1 to 2, by increments of 0.1.

2) The parameter β

Together with the size of the increment h, the value of the damping force parameter β was found to be most crucial in determining the stability of both the stochastic system and its deterministic counterpart. The influence of β on the stability of the deterministic system was described above. The stochastic system exhibited a similar behavior for all α (> 0) values used (see Figures 1 and 2 hereafter). In preliminary simulations, the stochastic system was found unstable for $\beta \leq 0$ and, therefore, negative β values were deemed less interesting to investigate. On the other hand, due to the upper bound condition imposed on h, large β values also forced the value of the h increment to decrease. Since, as explained above, this was not suitable, β was first taken to range from 0.00 to 1.50, by increments of 0.01.

Moreover, in order to better describe the "crossover" pattern exhibited for large β values (see Figure 3), another simulation run was performed for β values ranging from 1.00 to 2.00 (using a lower h value).

3) The parameter h

The value of the discrete time increment h had to be chosen carefully. Indeed, varying the h value (which the simulation program will do if h does not satisfy the

upper bound condition derived in Subsection 6.2) yields deterministic systems with different stability features. Simulation results showed that the same is true for the stochastic system. For very small h values and when β is small, the discrepancies between the undisturbed (deterministic) system and the stochastic systems (even though present) were less striking. On the other hand, the choices of α and β values were partly determined by considerations about the h value and conversely. Hence, with α ranging from 1.00 to 2.00 and β ranging from 0.00 to 1.50, the value h = 0.35 was found to be appropriate for this simulation. As h decreases to zero, it would be quite interesting to investigate the convergence of the Lyapunov exponents of the discretized stochastic systems towards (possibly) their continuous counterparts. Finally, when investigating the "crossover" pattern exhibited by the Lyapunov exponent curves for large β values, the step size h was decreased to 0.20.

The simulation results are given in Tables 1 and 2, as well as in Figures 1 through 3 hereafter.

Table 1 reports selected values of the top Lyapunov exponent for various combinations of α and β (h = 0.35). Recall that the case $\alpha = 0$ corresponds to the deterministic system, in which case λ is simply the logarithmic norm of the largest eigenvalue for the matrix A (h). This table shows that there are (α, β) combinations for which the stochastic system turns out to be unstable ($\lambda \geq 0$), while the corresponding deterministic system was stable ($\alpha = 0$ and same β value). As α increases, this disparity in the stability behavior of the stochastic and deterministic systems becomes more obvious. This illustrates the fact that attempts to deal with

random systems by averaging the noise (which in this case simply yields the deterministic system) are not reliable.

Figure 1 reports curves of (λ, β) values for several α's. Besides the already mentioned fact that the stochastic systems ($\alpha > 0$) do exhibit instability for some β's (when α becomes sufficiently large), while the corresponding deterministic system ($\alpha = 0$) is stable, it can also be seen that, for small β values, the rate of exponential decrease was always slower for the stochastic systems than for the corresponding deterministic systems. On the other hand, when, for larger β values, the deterministic system becomes overdamped, yielding increasing λ values, the stochastic systems seemed to "react more slowly" to this overdamping, i.e.:

1) The stochastic systems reached their minimal λ value for larger β values than the deterministic system and, as α increased, the β values at which the stochastic systems reached their fastest rate of exponential stability tended to increase.

2) Overdamping, which causes the deterministic system to exhibit a rapid decrease in its rate of exponential decay, seemed to influence the stochastic systems in a less marked way, causing the (λ, β) curves for the stochastic and deterministic systems to cross.

The "crossover" behavior, for large β values, of the (λ, β) curves is most obvious from the data in Table 2 and in Figure 2, whose purpose was specifically to investigate the range of β values at which this behavior was obseved (using a lower h value of 0.20). The displacement of the minima of the (λ, β) curves and their "crossover" behavior was also observed in the continuous time case (see Arnold and Kliemann (1983, Section IV C)).

Finally, Figure 3 gives a three dimensional representation of the stability behavior of the simulated stochastic systems. Again note that the (α, β) region in which these systems are stable corresponds to the cleft in the three dimensional plot ($\lambda < 0$). As α increases, this cleft (which crosses the range of α values) tends to get narrower (in terms of β values) and shallower, again indicating that increased α values resulted, for lower β values, in more instability (less stability) for the stochastic systems.

Table 6.1. Simulated top Lyapunov exponents for selected combinations of α and β [a, b]

$\beta \rightarrow$ $\alpha \downarrow$	0.0	0.2	0.4	0.6	0.8	1.0	1.2	1.4
0.0	0.06	–0.01	–0.09	–0.18	–0.29	–0.43	–0.21	–0.16
1.0	0.06	0.00	–0.08	–0.15	–0.21	–0.26	–0.27	–0.19
1.1	0.09	0.02	–0.05	–0.12	–0.20	–0.27	–0.26	–0.19
1.2	0.09	0.02	–0.04	–0.11	–0.19	–0.26	–0.27	
1.3	0.07	0.00	–0.07	–0.12	–0.18	–0.24	–0.25	
1.4	0.07	0.01	–0.07	–0.12	–0.18	–0.23	–0.25	
1.5	0.08	0.01	–0.06	–0.11	–0.17	–0.21	–0.24	
1.6	0.09	0.01	–0.06	–0.11	–0.16	–0.21	–0.24	
1.7	0.10	0.01	–0.05	–0.10	–0.16	–0.20	–0.24	
1.8	0.10	0.02	–0.05	–0.10	–0.15	–0.20	–0.22	
1.9	0.10	0.02	–0.04	–0.09	–0.15	–0.19	–0.22	
2.0	0.11	0.03	–0.04	–0.09	–0.14	–0.18	–0.22	

[a] The step size is h = 0.35. Missing entries correspond to those (α, β) combinations for which h = 0.35 was not acceptable.

[b] Recall that $\alpha = 0.0$ corresponds to the deterministic difference equation $x_{n+1} = A(h)\, x_n$.

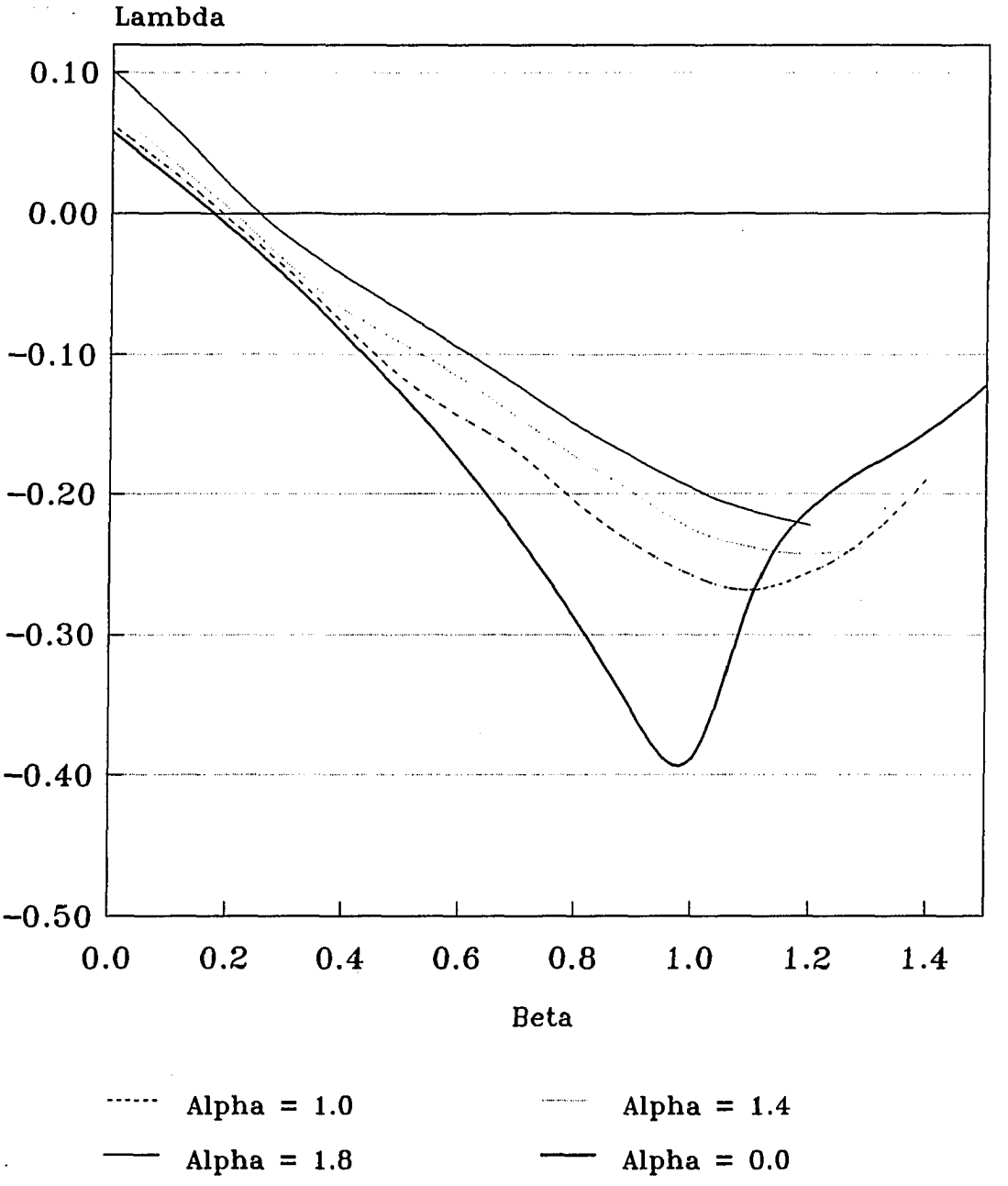


Figure 6.1. **Curves for the variation of λ as a function of β at selected α values (h = 0.35)**

Table 6.2. The values of λ for large β values (crossover region) [a]

$\beta \rightarrow$ $\alpha \downarrow$	1.0	1.2	1.4	1.6	1.8	2.0
0.0	–0.22	–0.11	–0.09	–0.07	–0.06	–0.05
1.0	–0.15	–0.15	–0.10	–0.08	–0.07	–0.06
1.1	–0.15	–0.15	–0.10	–0.08	–0.07	–0.06
1.2	–0.15	–0.14	–0.10	–0.08	–0.07	–0.06
1.3	–0.14	–0.15	–0.11	–0.08	–0.07	–0.06
1.4	–0.14	–0.14	–0.11	–0.08	–0.07	–0.06
1.5	–0.13	–0.14	–0.11	–0.08	–0.07	–0.06
1.6	–0.13	–0.14	–0.12	–0.09	–0.07	–0.06
1.7	–0.13	–0.14	–0.13	–0.09	–0.07	–0.06
1.8	–0.13	–0.14	–0.13	–0.09	–0.07	–0.06
1.9	–0.12	–0.13	–0.13	–0.09	–0.07	–0.06
2.0	–0.12	–0.13	–0.13	–0.10	–0.08	–0.06

[a] The step size is h = 0.20.

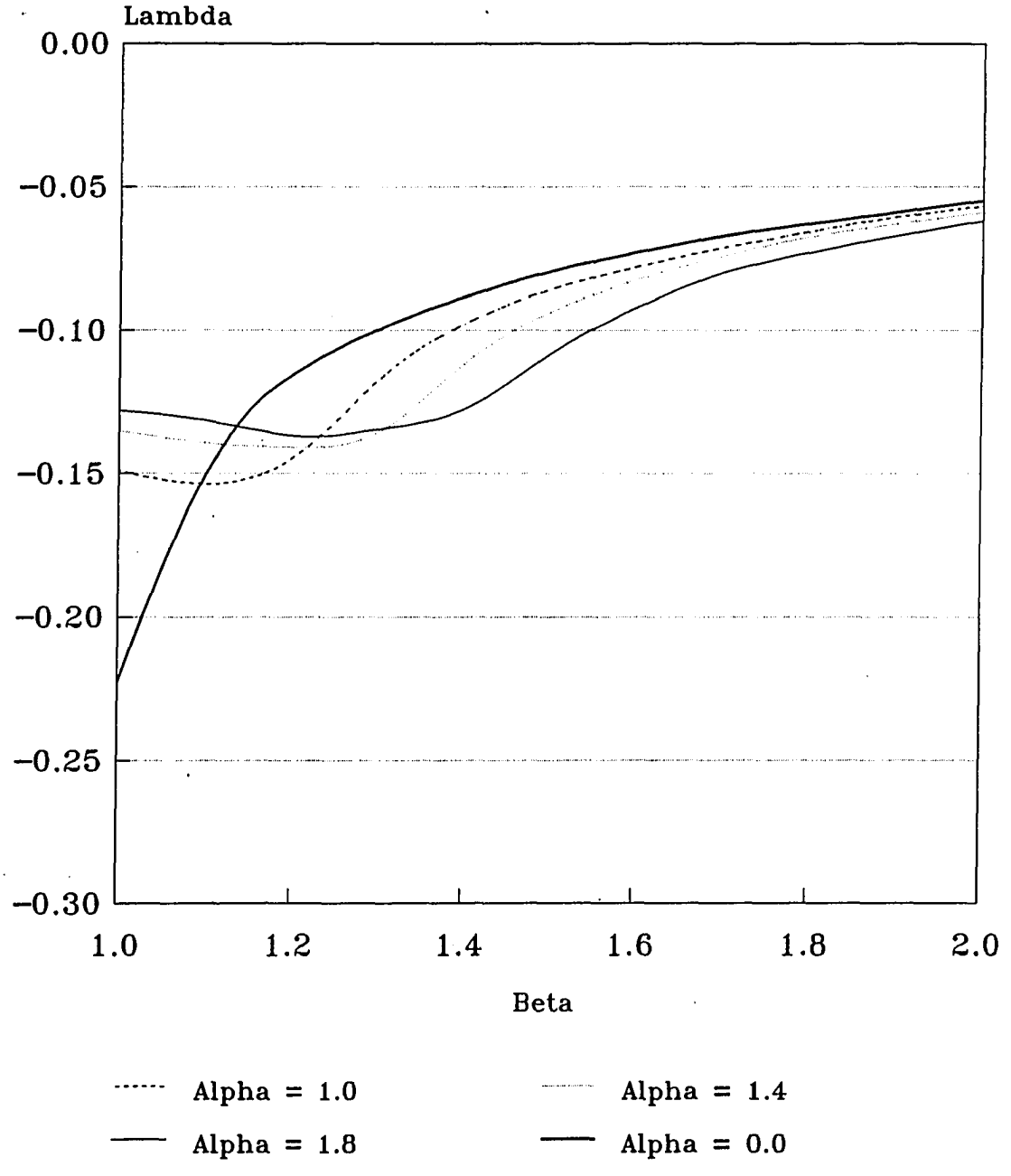


Figure 6.2. Curves for the variation of λ as a function of β at selected α values in the crossover region (h = 0.20)

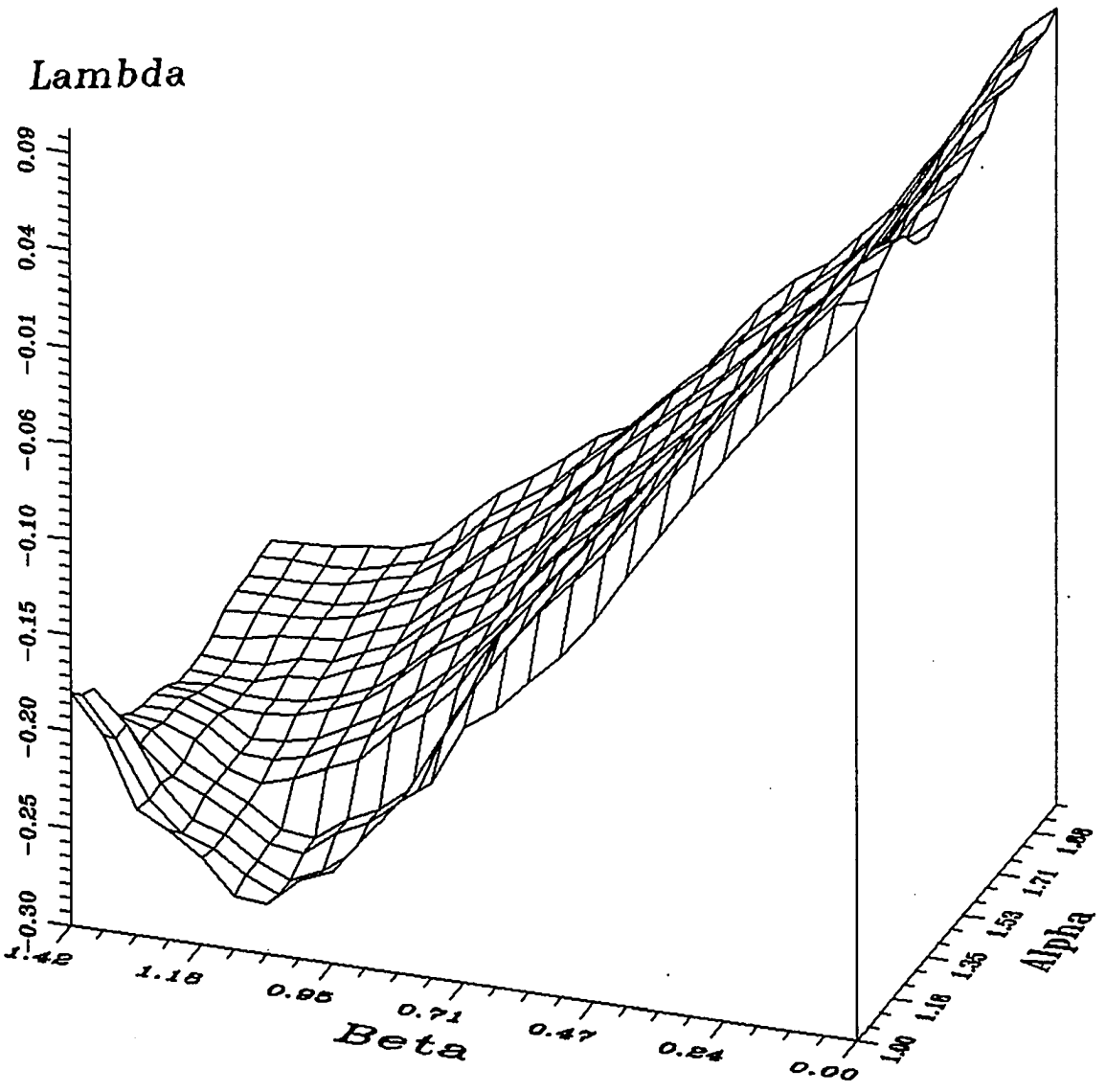


Figure 6.3. Three dimensional plot of the λ values for all the (α, β) combinations used in the main simulation (h = 0.35)

7. BIBLIOGRAPHY

Arnold, L. (1984). A Formula Connecting Sample and Moment Stability of Linear Stochastic Systems. *Siam Journal of Applied Mathematics* **44**, No 4 : 793–802.

Arnold, L. and Kliemann, W. (1983). Qualitative Theory of Stochastic Systems. Pp. 1–79 in Probabilistic Analysis and Related Topics, Vol **3** (A. T. Bharucha–Reid editor). Academic Press, New York.

Arnold, L. and Kliemann, W. (1986). Large Deviations of Linear Stochastic Differential Equations. Pp. 1–35 in Proceedings of the Conference on Stochastic Differential Systems in Eisenach, GDR (H. J. Engelbert and W. Schmidt editors). Springer Lecture Notes in Control and Information Sciences **96**. Springer–Verlag, Berlin, New York.

Arnold, L. and Kliemann, W. (1987). On Unique Ergodicity for Degenerate Diffusions. *Stochastics* **21** : 41–61.

Arnold, L., Kliemann, W., and Oeljeklaus, E. (1986a). Lyapunov Exponents of Linear Stochastic Systems. Pp. 85–128 in Lyapunov Exponents (L. Arnold and V. Wihstutz editors). Lecture Notes in Mathematics **1186**. Springer–Verlag, New York.

Arnold, L., Oeljeklaus, E., and Pardoux, E. (1986b). Almost Sure and Moment Stability for Linear Ito Equations. Pp 129–159 in Lyapunov Exponents (L. Arnold and V. Wihstutz editors). Lecture Notes in Mathematics **1186**. Springer–Verlag, New York.

Arnold, L. and Wihstutz, V. (1986). Lyapunov Exponents : A Survey. Pp. 1–26 in Lyapunov Exponents (L. Arnold and V. Wihstutz editors). Lecture Notes in Mathematics **1186**. Springer–Verlag, New York.

Boothby, W. M. (1986). An Introduction to Differentiable Manifolds and Riemannian Geometry. Academic Press, New York.

Boothby, W. M. and Wilson, E. N. (1979). Determination of the Transitivity of Bilinear Systems. *SIAM Journal on Control and Optimization* **17**, No 2 : 212–221.

Bougerol, P. (1985). Théorèmes de la Limite Centrale pour les Produits de Matrices en Dépendance Markovienne. Résultats récents. Pp. 225–240 in Probability Measures on Groups, VIII (Oberwolfach). Lecture Notes in Mathematics **1210**. Springer–Verlag, Berlin, New–York.

Bougerol, P. (1986a). Comparaison des Exposants de Lyapounov des Processus Markoviens Multiplicatifs. Preprint.

Bougerol, P. (1986b). Limit Theorems for Products of Random Matrices with Markovian Dependence. To appear in the Proceedings of the First International Congress of the Bernoulli Society, Tachkent (1986).

Bougerol, P. (1987). Tightness of Products of Random Matrices and Stability of Linear Stochastic Systems. *The Annals of Probability* **15**, No 1 : 40–74.

Bougerol, P. (1988). Théorèmes Limite pour les Systèmes Linéaires à Coefficients Markoviens. *Probability Theory and Related Fields* **78** : 193–221.

Bougerol, P. and Lacroix, J. (1985). Products of Random Matrices with Applications to the Schrödinger Operator. Progress in Probability and Statistics **8**. Birkhäuser, Boston.

Breiman, L. (1968). Probability. Addison–Wesley, Reading.

Buck, R. C. (1978). Advanced Calculus. McGraw–Hill, New York.

Bunke, H. (1972). Gewöhnliche Differentialgleichungen mit Zufälligen Parametern. Mathematische Lehrbücher und Monographien, Band **31**. Akademie–Verlag, Berlin.

Chung, K. L. (1964). The General Theory of Markov Processes According to Doeblin. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **2** : 230–254.

Chung, K. L. (1974). A Course in Probability Theory. Probability and Mathematical Statistics **21**. Academic Press, Orlando.

Cogburn, R. (1975). A Uniform Theory for Sums of Markov Chain Transition Probabilities. *The Annals of Probability* **3**, No 2 : 191–214.

Colonius, F. and Kliemann, W. (1989). Infinite Time Optimal Control and Periodicity. *Applied Mathematics and Optimization* **20** : 113–130.

Crauel, H. (1987). Random Dynamical Systems : Positivity of Lyapunov Exponents, and Markov Systems. Ph. D. thesis, Bremen University (GDR).

Doob, J. L. (1953). Stochastic Processes. John Wiley, New York.

Ellis, R. S. (1985). Entropy, Large Deviations, and Statistical Mechanics. Springer–Verlag, New York.

Ethier, S. N. and Kurtz, T. G. (1986). Markov Processes – Characterization and Convergence. John Wiley, New York.

Furstenberg, H. (1963). Noncommuting Random Products. *Transactions of the American Mathematical Society* **108** : 377–428.

Furstenberg, H. and Kesten, H. (1960). Products of Random Matrices. *The Annals of Mathematical Statistics* **31** : 457–469.

Has'minskiĭ, R. Z. (1967). Necessary and Sufficient Conditions for the Asymptotic Stability of Linear Stochastic Systems. *Theory of Probability and its Applications* **12** : 144–147.

Has'minskiĭ, R. Z. (1980). Stochastic Stability of Differential Equations. Sijthoff and Noordhoff, Alphen aan den Rijn. (Translation of the Russian edition, Moscow, Nauka (1969).)

Helgason, S. (1978). Differential Geometry, Lie Groups, and Symmetric Spaces. Academic Press, New York.

Hewitt, E. and Stromberg, K. (1965). Real and Abstract Analysis. Springer–Verlag, New York.

Isidori, A. (1985). Nonlinear Control Systems : An Introduction. Springer–Verlag, Berlin, New York.

Jain, N. and Jamison, B. (1967). Contribution to Doeblin's Theory of Markov Processes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **8** : 19–40.

Jakubczyk, B. and Sontag E. D. (1988). Controllability of Nonlinear Discrete Time Systems : A Lie–Algebraic Approach. To appear in *Siam Journal on Control and Optimization.*

Kliemann, W. (1979). Qualitative Theorie NichtLinearer Stochastischer Systeme. Ph.D. thesis, Bremen University (GDR).

Kliemann, W. (1987). Recurrence and Invariant Measures for Degenerate Diffusions. *The Annals of Probability* **15**, No 2 : 690–707.

Kliemann, W. (1988). Analysis of Nonlinear Stochastic Systems. Pp. 43–102 in Analysis and Estimation of Stochastic Mechanical Systems (Udine 1987). CISM Courses and Lectures **303**. Springer–Verlag, Vienna.

Lyapunov, A. M. (1949). Problème Général de la Stabilité du Mouvement. Annals of Mathematics Studies **17**. Princeton University Press, Princeton, New Jersey. (Reprint).

MacDonald, I. G. (1979). Algebraic Structure of Lie Groups. London Mathematical Society, Lecture Notes Series **34** : 91–150.

Mañé, R. (1987). Ergodic Theory and Differentiable Dynamics. Ergebnisse der Mathematik und ihrer Grenzgebiete 3. Folge – Band 8. Springer–Verlag, Berlin, Heidelberg.

Meyn, S. P. (1989). Ergodic Theory and Topological Dynamics for Controlled Discrete Time Systems. Preprint.

Meyn, S. P. and Caines, P. E. (1988). Asymptotic Behavior of Stochastic Systems Possessing Markovian Realizations. Preprint.

Miller, R. K. and Michel, A. N. (1982). Ordinary Differential Equations. Academic Press, New York.

Munkres, J. R. (1975). Topology, a First Course. Prentice–Hall, Englewood Cliffs, New Jersey.

Orey, S. (1971). Lecture Notes on Limit Theorems for Markov Chain Transition Probabilities. Van Nostrand Reinhold, New York.

Oseledeč, V. I. (1968). A Multiplicative Ergodic Theorem. Lyapunov Characteristic Numbers for Dynamical Systems. *Transactions of the Moscow Mathematical Society* **19** : 197–231.

Revuz, D. (1984). Markov Chains. North Holland Mathematics Library, Elsevier Science Publisher, New York.

Rockafellar, T. R. (1970). Convex Analysis. Princeton University Press. Princeton, New Jersey.

San Martin, L. and Arnold, L. (1986). A Control Problem Related to the Lyapunov Spectrum of Stochastic Flows. *Matemática Applicada e Computacional* **5**, No. 1 : 31–64.

Stettner, L. (1988). On Ergodic Decomposition of Feller Markov Processes. To appear.

Tweedie, R. L. (1975). Sufficient Conditions for Ergodicity and Recurrence of Markov Chains on a General State Space. *Stochastic Processes and their Applications* **3** : 385–403.

Tweedie, R. L. (1976). Criteria for Classifying General Markov Chains. *Advances in Applied Probability* **8** : 737–771.

8. ACKNOWLEDGEMENT

The author wishes to express his gratitude to Dr. Krishna Athreya. The pleasure we had in interacting with him for so many years and the interest generated by his teaching are the reason why we became interested in probability theory and stochastics.

Many thanks also to our committee members : Drs. Dean Isaacson, Ken Koehler, Yasuo Amemiya, and Elgin Johnston. All of them were helpful at all times and contributed a great deal to the completion of the degree whose requirements are partly fulfilled through this work.

And then, there is Dr. Wolfgang Kliemann. He directed this work over the last five years and most of what we know in this field, we learned through endless discussions and the infinitely many (P–a.s.) questions he relentlessly answered. He offered countless suggestions for improvement, helped us tighten arguments, revise proofs, etc. He supported us literally over land and sea. Without him this thesis would not be.

9. APPENDIX: THE SIMULATION PROGRAM

9.1. The Simulation Program

The program used for the simulation discussed in Section 6 was written in Turbo Pascal (Version 4). It was designed in a very modular manner and consists of a main program (**Lya.pas**), which basically only handles the loops used to go through all the α and β combinations requested, and then calls procedures in several Turbo Pascal units.

The Turbo Pascal units called by **Lya.pas** (directly or via another unit) are:

1) **Dos** and **Crt** (two units included with the Turbo Pascal package).
2) **MatUnit** (containing the procedures and functions needed for the matrix manipulations in the program).
3) **MathUnit** (containing various mathematical procedures and functions).
4) **ScrnUnit** (containing procedures and functions related to screen output).
5) **FileUnit** (containing procedures and function related to file output).
6) **SimUnit** (containing all the procedures and functions written specifically for this simulation).

The length of the entire program precludes its inclusion in full in this appendix. Nevertheless, a few of the procedures, which are at the heart of the simulation, are

given (even though somewhat abridged for the sake of clarity) in the next few pages. These procedures are not self–standing since they themselves call several unlisted procedures or functions, which are only briefly described by comments in the listing.

Among the procedures which we elected not to list hereafter, but which deserve a few comments, are **Scaling** and **Get_H_Increment**.

The entries of the matrix product $A(\xi_n) \ldots A(\xi_0)$ will very often become extremely large (over the order E+300). **Scaling** is a procedure which periodically divides the entries of the matrix product $A(\xi_n) \ldots A(\xi_0)$ by the absolute value of its largest entry and keeps track of these divisors to correct the final answer (i.e., the simulated top Lyapunov exponent). This is necessary to prevent the program from crashing when attempting to compute the (Euclidean) norm of these matrices. Indeed, the computation of such a norm first involves summing the square of the matrix entries and, without rescaling, this intermediate step would generate numbers of an unmanageable magnitude.

The procedure **Get_H_Increment** checks whether the initially selected time increment h does, for the selected (α, β) combinations, satisfy the upper bound condition discussed in Subsection 6.2. If it does not, the procedure will successively reduce the value of h by decrements of 0.5, until the new h satisfies the given upper bound condition. This procedure is called in the main program **Lya.pas**, before any attempt is made to obtain the top Lyapunov exponent for any (α, β) combination.

The main part of the simulation is called via the procedure **Lyapunov**. This procedure is called in the main program for each (α, β) combination and, relying on calls to many other procedures or functions, computes the top Lyapunov exponent for the stochastic system corresponding to the chosen (α, β) combination. The procedure **Lyapunov**, together with a few other key procedures, is listed hereafter.

```
Unit Simunit ;    {Contains the Procedures and Functions specific}
                  {to the simulation program.                    }

(*************************************************************************)
Interface
(*************************************************************************)

(*$N+*)                    {Math Coprocessor (Double precision)}
                          {compiler directive.                 }

Uses Dos,Crt,MatUnit,MathUnit,ScrnUnit ;

        {Define Matrix and FnArray as Array[1..MatSize] and   }
        {Array[1..IterationSize] of Double,  where MatSize and}
        {IterationSize are fixed constants.                   }

        {MatSize is declared and set to 4 in the unit MatUnit.}

Type
      MatSequence = Array[1..IterationSize] of Matrix ;

                  {A variable of the type MatSequence }
                  {(Semigroup) will contain the last  }
                  {IterationSize number of products of}
                  {random matrices obtained,  i.e.,   }
                  {A(n+IterationSize)...A(0) up to    }
                  {A(n)...A(0).                       }

                  {If MatSize is too large,  this structure}
                  {will be too large.                      }

Var
      SequenceToLambda    : FnArray ;

                  {SequenceToLambda will contain the the last}
                  {IterationSize values for the log norm of  }
                  {the matrices in MatSequence (see above).  }

                  {SequenceToLambda is the tail of the sequence  }
                  {of values converging to the Lyapunov exponent.}

      FirstStepSemigroup : Matrix  ;

                  {FirstStepSemigroup is first A(0) and then, }
                  {since the sequences MatSequence and        }
                  {SequenceToLambda are handled in groups of  }
                  {size IterationSize, FirstStepSemigroup is  }
                  {reset to be the last element of MatSequence}
                  {from the previous iteration times a new    }
                  {random matrix A.                           }
```

```
      H,                                {H is the time increment. }

      Alpha,                            {That's the factor in the.}
                                        {restoring force.         }
      Beta              : Double  ;     {The damping force.       }

      N                 : Integer ;     {A counter.}

{-------------------------------------------------------------------}

Procedure GetNewMatNorm  (IterationSize,
                          N,
                          Step             : Integer ;
                          StepMatrix       : Matrix  ;
                      Var Correction       : Double  ;
                      Var SequenceToLambda : FnArray ) ;

{Procedure to specifically compute the terms of the FnArray}
{SEQUENCETOLAMBDA whose elements converge to lambda.       }

{-------------------------------------------------------------------}

Procedure LambdaSequence  (H,
                           Precision          : Double  ;
                           IterationSize      : Integer ;
                       Var N                  : Integer ;
                       Var FirstStepSemigroup : Matrix  ;
                       Var Correction         : Double  ;
                       Var SequenceToLambda   : FnArray ) ;

{This Procedure combines several other procedures to generate the}
{sequence SEQUENCETOLAMBDA) which converges to lambda.           }

{-------------------------------------------------------------------}

Procedure Lyapunov  (H,
                     Precision : Double  ;
                     LimitTest : Boolean ;
                 Var N         : Integer ;
                 Var Overflow  : Boolean ;
                 Var Limit     : Double  ) ;

{Procedure to actually obtain the top Lyapunov exponent as the limit}
{of the sequence of matrix norms in the SequenceToLambda array.     }

(*********************************************************************)
Implementation
(*********************************************************************)
```

```
 Procedure GetNewMatNorm ;

Var
   Norm    : Double      ;

   i       : Integer     ;

Begin

   MatNorm(4,StepMatrix,Norm) ;  {Procedure to Compute the Euclidean}
                                 {norm of the matrix StepMatrix. The}
                                 {norm is outputed in the variable  }
                                 {Norm.                             }

   {The next step actually computes 1/n log |A(n)...A(0)| when}
   {n=IterationSize*N + Step. The variable Correction is used }
   {to undo the rescaling performed by the procedure SCALING. }

   SequenceToLambda[Step] := (Ln(Norm) + Correction)
                             / (IterationSize*N + Step) ;

End ;

{--------------------------------------------------------------------}

Procedure LambdaSequence ;

Var
   ZeroMatrix       : Matrix      ;

   Semigroup        : MatSequence ;

   Sequence,
   NormSemigroup    : FnArray     ;

   i                : Integer     ;

   Norm             : Double      ;

Begin

   {The following loops initialize variables to zero.}

   For i := 1 to 4 Do
      ZeroMatrix[i] := 0 ;

   For i := 1 To IterationSize Do
      Semigroup[i] := Zeromatrix ;
```

```
      For i := 1 To IterationSize Do
        Begin
           NormSemigroup[i] := 0 ;
           SequenceToLambda[i] := 0 ;
        End ;

      {The first value of the matrix (sequence) Semigroup is set to}
      {a random matrix whose entries are in FirstStepSemigroup.     }

      Semigroup[1] := FirstStepSemigroup ;

      {The next procedure simply obtains the value of 1/n log |A(n)...A(0)|}
      {when n=IterationSize*N+1. The answer is passed to SequenceToLambda. }

      GetNewMatNorm(IterationSize,N,1,Semigroup[1],Correction,
                    SequenceToLambda) ;

      For i := 2 To IterationSize Do
         Begin

            {First generate A(IterationSize*N+i)...A(0)}
            {in Semigroup[i] through multiplication of }
            {A(IterationSize*N+(i-1)...A(0) in         }
            {Semigroup[i-1] by a new random matrix A.  }

            GetNewSemigroup(H,Semigroup[i-1],Semigroup[i]) ;

            {The next procedure simply obtains the value of}
            {1/n log |A(n)...A(0)| when                    }
            {n=IterationSize*N+i. The answer is passed to  }
            {SequenceToLambda.                             }

            GetNewMatNorm(IterationSize,N,i,Semigroup[i],Correction,
                          SequenceToLambda) ;

         End ;

      {For the next iteration (involving IterationSize elements), the}
      {new matrix FirstStepSemigroup is set to be the last element of}
      {the previous iteration, Semigroup[IterationSize], times a new }
      {random matrix. The next time the current procedure is invoked,}
      {Semigroup[1] is then set to FirstStepSemigroup (see above).   }

      GetNewSemigroup(H,Semigroup[IterationSize],FirstStepSemigroup) ;

      N := N+1 ;     {Update the counter for the number of times the  }
                     {program had to check the limit based on the last}
                     {IterationSize elements of SequenceToLambda.     }

End ;
```

```
{-------------------------------------------------------------------------}

Procedure Lyapunov ;

Var
   b                : Byte    ;

   i                : Integer ;

   Correction       : Double  ;

Begin

   {Initialize the first matrix by computing a random matrix, stored}
   {in OneStepSemigroup, and assigning this initial random matrix to}
   {FirstStepSemigroup for use in the procedure LAMBDASEQUENCE.     }

   SetNewMatrix(H,OneStepSemiGroup) ;

   FirstStepSemigroup := OneStepSemigroup ;


   {Initialize the Boolean variable indicating that a limit was found}
   {to False, set the counter N (= number of times the LambdaSequence}
   {procedure was used) to 0, initialize the correction variable to  }
   {False.                                                           }

   LimitTest        := False ;
   Correction       := 0     ;
   N                := 0     ;

   {Compute the limit (lambda) by generating a sequence of values (using}
   {the LambdaSequence procedure repeatedly) until the CheckForLimit    }
   {procedure yields a limit (LimitTest becomes True).                  }

   Repeat

      {Since the entries of the matrix product                        }
      {A(IterationSize*N+1)...A(0) (stored in FirstStepSemigroup)}
      {can become very large, the next procedure rescales these  }
      {entries by dividing each of them by the largest absolute  }
      {value for all the entries. The resulting matrix is stored }
      {in (a new) FirstStepSemigroup and the correction needed to}
      {obtain the exact value for 1/n log |A(n)...A(0)| without  }
      {this rescaling is stored in the variable Correction.      }

      Scaling(FirstStepSemigroup,FirstStepSemigroup,Correction) ;

      {The next procedure actually generates the sequence of values}
      {1/n log |A(n)...A(0)| by groups of IterationSize values.    }
```

```
        LambdaSequence(H,Precision,IterationSize,N,FirstStepSemigroup,
                       Correction,SequenceToLambda) ;

      If (N*IterationSize > 32500)      {This If statement simply checks}
         Then                           {that the number of total       }
           Begin                        {iterations to obtain the limit }
              Overflow := True ;        {lambda (in the variable Limit) }
              Exit ;                    {does not exceed 32500. If yes, }
           End ;                        {the simulation quits.          }

      {The following procedure uses the last IterationSize elements of}
      {the sequence SequenceToLambda to see if a limit was attained at}
      {the selected precision. If yes, LimitTest is set to True and   }
      {the limit (Lambda) is stored in the variable Limit.            }

      CheckForLimit(SequenceToLambda,Precision,IterationSize,
                    LimitTest,Limit) ;

   Until (LimitTest = True) ;

End ;

(*************************************************************************)

End.
```